DOCUMENT RESUME

ED 239 558 HE 016 982

AUTHOR	Friedman, Miriam; And Others
TITLE	Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Social Interaction Generic Instrument. Final Report, Research Report Number Two.
INSTITUTION	Alverno Coll., Milwaukee, Wis.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	82
GRANT	NIE-G-77-0058
NOTE	91p.; For related documents, see HE 016 980-990.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	College Curriculum; College Students; Competence; *Competency Based Education; Evaluation Methods; Higher Education; *Interpersonal Competence; Leadership; *Liberal Arts; Outcomes of Education; *Research Methodology; *Social Behavior; *Student Evaluation; Test Validity
IDENTIFIERS	*Alverno College WI

ABSTRACT

The methodology for validating assessment techniques in a performance-based liberal arts curriculum was studied at Alverno College. Methods for validating a generic competence instrument that measures "social interaction" were investigated. A generic instrument is defined as one that assesses a competence level across content areas instead of using a large variety of instruments. The assessment system at the college requires students to demonstrate incremental gains while progressing through sequential levels in each competency area. A comparison was undertaken of the performance of 69 students on entrance to college with that of 32 students who had 2 years of college instruction on each of the social interaction dimensions (preparation, demonstration, self-assessment, and leadership). Including students with a broad range of age and formal learning experience led to an effective strategy for identifying those social interaction behaviors that validate the construct. The analysis indicated that the study of assessment techniques should not be limited to univariate methods. Attention is directed to instructional validity, construct validity, instrument criterion validity, and a leadership dimension (the relationship between leadership and type of program, age, and other college experience). (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED239558

VALIDATING ASSESSMENT TECHNIQUES IN AN OUTCOME–CENTERED LIBERAL ARTS CURRICULUM: SOCIAL INTERACTION GENERIC INSTRUMENT

Miriam Friedman Marcia Mentkowski
Bernardin Deutsch M. Nicolette Shovar Zita Allen

Office of Research & Evaluation/Social Interaction Division
ALVERNO COLLEGE

FINAL REPORT TO THE NATIONAL INSTITUTE OF EDUCATION:
RESEARCH REPORT NUMBER TWO

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Funded by a grant from the National Institute of Education:
Careering After College: Establishing the Validity of Abilities Learned in College for Later Success
(NIE–G–77–0058)

Principal Investigators:
Marcia Mentkowski
Austin Doherty
Alverno College
3401 South 39th Street
Milwaukee, Wisconsin 53215

HE 016 982

An overview and rationale for our approach to the study of college outcomes, and a summary of the results from the following series of ten research reports, are found in:

Marcia Mentkowski and Austin Doherty. Careering After College: Establishing the Validity of Abilities Learned in College for Later Careering and Professional Performance. Final Report to the National Institute of Education: Overview and Summary. Milwaukee, WI: Alverno Productions, 1983.

Research Reports:

One: Friedman, M., Mentkowski, M., Earley, M., Loacker, G., & Diez, M. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Valuing and Communications Generic Instrument, 1980.

Two: Friedman, M., Mentkowski, M., Deutsch, B., Shovar, M.N., & Allen, Z. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Social Interaction Generic Instrument, 1982.

Three: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Insights From the Evaluation and Revision Process, 1980.

Four: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Integrated Competence Seminar, 1982.

Five: Assessment Committee/Office of Research and Evaluation. Validating Assessment Techniques in an Outcome-Centered Liberal Arts Curriculum: Six Performance Characteristics Rating, 1983.

Six: Mentkowski, M., & Strait, M. A Longitudinal Study of Student Change in Cognitive Development and Generic Abilities in an Outcome-Centered Liberal Arts Curriculum, 1983.

Seven: Much, N., & Mentkowski, M. Student Perspectives on Liberal Learning at Alverno College: Justifying Learning as Relevant to Performance in Personal and Professional Roles, 1982.

Eight: Mentkowski, M., Much, N., & Giencke-Holl, L. Careering After College: Perspectives on Lifelong Learning and Career Development, 1983.

Nine: Mentkowski, M., DeBack, V., Bishop, J., Allen, Z., & Blanton, B. Developing a Professional Competence Model for Nursing Education, 1980.

Ten: Mentkowski, M., O'Brien, K., McEachern, W., & Fowler, D. Developing a Professional Competence Model for Management Education, 1982.

© Copyright 1982. Alverno College Productions, Milwaukee, Wisconsin. All rights reserved under U.S., International and Universal Copyright Conventions. Reproduction in part or whole by any method is prohibited by law.

ABSTRACT

This report explores issues related to the validation of more nontraditional assessment techniques, and tests some ways such studies may proceed. We explore the appropriateness of various methods for validating a generic competence instrument that measures Social Interaction, a construct with little or no history as a teachable college outcome or measure. We compare the performance of 69 uninstructed students on entrance to college with that of 32 students who had two years of college instruction on each of the Social Interaction dimensions (Preparation, Demonstration, Self-Assessment and Leadership), and the specific dimension criteria.

Results indicate similarities in performance between tracitional age instructed students and mature uninstructed students. While this may be expected, it also indicates that group comparisons may not be an effective strategy for validating assessment techniques if the ability is one developed through prior informal as well as college learning. Despite our efforts to do so, we were not able to control for the myriad range of variables that are likely to affect the results. When performance of such an ability also interacts with a set of personal and ego development variables, separating out the specific effects of instruction that show significant differences through group comparison is not an effective strategy, especially given the small sample sizes generally available.

However, some Social Interaction criteria did indeed separate the uninstructed students from the instructed students when we combined all students in a discriminant analysis. These criteria are more closely related to those aspects of Social Interaction that are learned as part of the more specific Social Interaction learning experiences. Thus, including students with a broad range of age and formal learning experience did lead to an effective strategy for identifying those Social Interaction behaviors that validate the construct. Clearly, the study of assessment techniques should not be limited to univariate methods; patterns of coherent group performance provide us with a more holistic picture of performance, particularly of Social Interaction, not well understood and measured compared to some other abilities like communications.

The present study outlines a procedure by which the integration of information about competence construct, different group characteristics and criteria evaluation contribute to an information base for instructional development, re-evaluation of competence definitions and revision of instrument criteria which measure these behaviors. The study helps to illuminate a key question in approaching the validation of any faculty designed instrument measuring important but not well defined abilities new to higher education instruction: What strategies are appropriate given where this instrument and construct are in their current development?

ACKNOWLEDGEMENTS

We wish to acknowledge the contributions of the following persons:

Office of Research & Evaluation:
Eunice Monroe
Laura Giencke-Holl
Michael J. Strait

Social Interaction Division (1979–82):
Zita Allen
Barbara Blanton
Barry Burd
Katherine Couture
Bernardin Deutsch
Jane Halonen
Ruth Hoerig
Patricia Jensen
Nancy Maynard
Clare Novak
Maria Terese Patterson
Penelope Reed
Ann Schlaefer
M. Nicolette Shovar

TABLE OF CONTENTS

	Page
INTRODUCTION	1
METHOD	10
Sample	10
Instrument	10
Procedure	13
Social Interaction Generic Criteria	16
Procedure for Coding Student Performance	17
Design	19
Analysis	22
RESULTS	23
Instructional Validity	23
Instrument Criteria Validity	38
Construct Validity	47
Leadership Dimension Score Analysis	57
CONCLUSIONS	61
Instructional Validity	61
Construct Validity	64
Criteria Validity	69
DISCUSSION	72
REFERENCES	78

VALIDATING ASSESSMENT TECHNIQUES IN AN OUTCOME-CENTERED LIBERAL ARTS CURRICULUM: SOCIAL INTERACTION GENERIC INSTRUMENT

Miriam Friedman Marcia Mentkowski Bernardin Deutsch

M. Nicolette Shovar Zita Allen

Office of Research & Evaluation/Social Interaction Division

ALVERNO COLLEGE

INTRODUCTION

The purpose of this paper is to continue to explore issues related to the validation of Alverno's more nontraditional assessment techniques, and to empirically illustrate some ways in which such validation studies may proceed. Validation of assessment techniques is a cornerstone in establishing the validity of the abilities learned in college (Mentkowski & Doherty, 1977, 1983). This study of the Social Interaction competence is another example of our work toward this goal. Other studies focused on the Communications and Valuing competences (Friedman, Mentkowski, Earley, Loacker and Diez, 1980), the Integrated Competence Seminar (Assessment Committee/Office of Research and Evaluation, 1982), the Six Performance Characteristics Rating (Assessment Committee/Office of Research and Evaluation, 1983), and a range of other instruments (Assessment Committee/ Office of Research and Evaluation, 1980).

Several validation methods have emerged as ways to examine assessment techniques in our previous work. In the present study, we investigate the extent to which the instrument measures the effects of instruction by comparing instructed and uninstructed student performance on each of

the Social Interaction dimensions as well as the specific criteria for each dimension. The second and perhaps more important objective is to examine patterns of student performance to assist in understanding student social interaction behavior and the meaning of the Social Interaction competence as it is currently defined.

The present study examines these validation methods in relation to the Social Interaction Generic Criteria used to judge performance at Social Interaction level 4 which students generally complete at the midpoint of their college career. In contrast to previous studies of Communications and Valuing where we studied levels 1 through 4, the Social Interaction Generic Criteria assess level 4 only.

Thus, the purpose of the current study is twofold. First, we study the appropriateness of various methods for validating generic competence instruments that assess constructs with little or no previous history as college outcomes. Methods for validating such open-ended constructs are still under examination. Then we examine several study designs, such that constraints imposed by the educational setting are recognized. Validation methods are considered for use at Alverno depending on how they will work within the context of the curriculum and the setting in which the results are applied. We are examining strategies that can be used within a regularly scheduled assessment process. Second, we are interested in adding to the understanding of how Social Interaction as an ability develops during college, and how informal learning experiences prior to formal college learning contribute to performance.

Faculty observe differences in student performance over two years in college given the nature of the assessment process. That is, each consecutive assessment is a "pretest" for the next competence level.

Thus, an interesting question is, How does the performance of instructed students compare to uninstructed students entering Alverno with more extensive life experiences?

First, experienced women can be expected to have social interaction abilities they have developed outside of college. Their social interaction skills may have been developed informally, without the benefit of college instruction. Thus, one interesting question has to do with how social interaction develops without the benefit of formal education. If we understand how it develops through experience, then we can better develop objectives and learning strategies that build on informal learning. To what extent should we expect that social interaction abilities can be developed in college, especially in traditional-aged students? Do we expect that traditional-aged students will develop abilities that match those of the mature woman? Do we expect that abilities will be significantly more developed due to the effects of instruction? If ego development or maturity is a major factor in the development of social interaction skills, to what extent should we expect that social interaction skills can be developed in college after two years?

Since we had been successful in demonstrating that traditional-aged, instructed students in the Weekday College had performed significantly higher on Valuing and Communications level 4 criteria compared to experienced, uninstructed students just entering Alverno in the Weekend College, we felt we could expect that such differences would also be obtained for the Social Interaction competence. This was also expected because we consider Valuing to be an ability that is less likely to be developed in the nontraditional-aged student through informal learning. Yet, at level 4, the nontraditional-aged students in Weekend College had not performed at level 4 as well as students with instruction. It should

be noted, however, that this was true only of level 4, and not of levels 1 to 3. Thus, nontraditional-aged students do develop some Valuing abilities through informal learning. In our Communications study, we found that the uninstructed group demonstrated a different developmental sequence in Communications competence, but with perfect scalability. We have reason, then, to believe that nontraditional-aged students do develop some of the competences through informal learning.

In the present study, we compare instructed students to uninstructed students. How might we best design such a study? This question is particularly important because our first purpose is to examine validation methods that can be carried out within a particular curriculum design. For the Social Interaction competence, one of several constraints in carrying out a "true" pre-/posttest design is that students at level 4 are required to complete five Social Interaction assessments.

The Social Interaction Generic Criteria are used to assess a student's demonstration of her ability on five separate occasions in order for her to receive a level 4 validation (credential). In order to examine student performance at level 4, one of these occasions must be chosen for a pre-/posttest design, one that is "content fair" (i.e., students may perform with or without a specific content background demanded by the instrument stimulus).

One assessment occasion that is content-fair on which uninstructed and instructed students may be compared, is the Integrated Competence Seminar Group Discussion (Assessment Committee/Office of Research and Evaluation, 1982b). This instrument is content-fair, and would be likely to produce results generalizable across the college.[1] (Note: In

[1]With many measures, it is difficult to give a pre-post assessment without the student having some pre-training in vocabulary. Without that, we may be testing just for vocabulary differences.

the Communications study, we controlled for content by selecting uninstructed students who had prior background in a subject area.) Since the instrument is content-fair, we can administer it as a pretest to uninstructed students. Currently, four of five Social Interaction instruments are content bound. The fifth is an independent study used as an external assessment. There are not enough students to provide us with a large enough sample size on this instrument and it would also be difficult to give as a pre-assessment.

Further, the ICS Group Discussion is currently the only social interaction performance that is videotaped, and requires written as well as oral performance. In the present study, we felt it important not to elicit the Hawthorne effect, that is, that students would be aware they were being observed as part of an instrument validation study, as well as for credentialing.

It was important to obtain a large enough sample size to yield statistically meaningful comparisons. In our first study of the Valuing competence, we had been able to obtain performance data on only 11 students who had contracted for validation at level 4 for comparison with students completing the pretest. Choosing the ICS Group Discussion as the instrument for the study guaranteed at least 30 students in the instructed student group.

In sum, the ICS Group Discussion seemed the most appropriate instrument for the study. It is content-fair, records of instructed student performance were on videotape, we had accumulated enough records to give us a large enough sample size, and records included written as well as videotaped oral performances.

How does instruction affect performance depending on the context in which the skills are demonstrated? Students demonstrating social interactions skills, along with other skills in a setting where they are not being credentialed, is perhaps the best test of their ability to transfer their social interaction ability. Demonstrating their ability in a more focused setting should be more likely to elicit their skills, especially if their performance is being counted toward a competence validation or credential.

Using the Integrated Competence Seminar Group Discussion meant that instructed students' performance would be measured in the context of the remaining parts of the ICS instrument, In-Basket and Oral Presentation. The instructed student is expected to demonstrate level 4 of Social Interaction at the same time that she is being assessed for level 4 of other competences as well. Thus, the instructed student would be challenged to perform a wide range of abilities, not just Social Interaction. It seemed inadvisable to ask entering students to perform all three parts of the ICS, since the rationale for their completing the pretest was for a level 1 validation in Social Interaction. Thus, the outcome is biased against the instructed students, rather than the uninstructed students. Still another aspect of the ICS is its use in the college for diagnostic purposes rather than for credentialing. Only some students were using the ICS Group Discussion to complete a level 4 validation. Thus, not only were we asking instructed students to complete the Group Discussion in the context of the complete Integrated Competence Seminar, we were also asking that they transfer their social interaction abilities beyond the classroom setting to an external situation where they were not being credentialed.

Since we had demonstrated that instructed students at level 4 performed significantly higher in Communications and Valuing at level 4 than uninstructed students, we felt we could make the assumption that level 4 instructed students would perform higher on a level 4 Social Interaction assessment that was not asking them to integrate this ability with others. Instructors had already assessed these students. Each student had successfully consecutively mastered levels 1 to 3. Given this assumption, how would students perform on the ICS Group Discussion in a situation where they had to integrate and transfer their abilities, and where some students were not being credentialed? We are also interested in how instructed students compare with students of nontraditional age who had some informal learning in social interaction skills.

Demonstrating that students can perform social interaction skills after instruction is not the most important question, however. Instructors report such increases in ability, and indeed, the fact that students are validated on successive levels of Social Interaction indicates growth. We wished to make a comparison that would pit older women against younger women in a setting which would call for transfer of learning on the part of the younger, instructed women. On the face of it, this meant a test of the differences in social interaction abilities between women who had experienced instruction in the Weekday College, and uninstructed women who were just entering the Weekend College.

The most important objective is the construct validity of the social interaction competence. Thus, it is more important that the group studied be a wide range of students, with and without experience. At the same time, however, we are interested in examining the differences between instructed and uninstructed students, keeping in mind that the

uninstructed groups will be given some prior instruction in vocabulary. Generally, our questions are:

- To what extent will instructed students differ in their social interaction performance?
- Do they conceptualize social interaction differently as compared with uninstructed students?
- What patterns of student performance characterize these students?

These questions, more specifically stated in collaboration with the Social Interaction Division and the Office of Research and Evaluation, served as a guide for the study design:

- How does instructed students' performance compare with uninstructed students' performance on each of the four Social Interaction dimensions of Preparation, Demonstration, Self-Assessment and Leadership (instructional validity)?
- How does instructed students' performance compare to uninstructed students on each of the behavioral criteria that comprise the dimensions (instrument criteria validity)?
  - To what extent do instrument criteria measure effects of instruction?
  - Which criteria best discriminate instructed from uninstructed groups?
  - Are all important social interaction skills at level 4 represented by the instrument's behavioral criteria?
  - Which criteria contribute most to effective Preparation, Demonstration and Self-Assessment?
- How might we better understand the meaning of the Social Interaction competence from an analysis of students' performance irrespective of differences between instructed and uninstructed students (construct validity)?
  - Can we identify patterns of Social Interaction behaviors which characterize each group of students? Does the Social Interaction construct, as a defined set of abilities, have the same meaning for different groups of students? Are

different groups characterized by different social interaction skills?

- As we review patterns of performance, how can we best describe the entry level of Social Interaction competence among uninstructed students who may have had informal learning experiences?
- How do patterns of performance in students who are uninstructed compare to students who have been instructed?
- What relationships exist between Leadership and the variables of type of program, instruction, age and other college experience?

METHOD

Sample

Sixty-nine uninstructed Weekend College women students, with no prior Alverno College experience, and 32 instructed Weekday College women students, with at least two years of Alverno College experience, comprise the sample of students selected for the study. Students ranged in age from 18 to 54 years with a mean age of 32 years. Students varied on the amount of prior college experience as well. They ranged from no prior college experience before entering Alverno to having achieved 120 credits of prior college experience before attending Alverno.

The 69 Weekend uninstructed students were drawn from all students entering Alverno Semester II, January 1980 (n = 110). Seventeen of the 110 students were re-entry students and were not included in the sample. Four students had previous Alverno College experience in the Weekday College; 10 did not participate for various scheduling reasons; 10 completed the instrument but their performance was not usable for the comparison study because assessors did not complete all the forms or because the recorder did not work. The remaining students (n = 69) comprised the final sample of uninstructed students.

The Weekday instructed group included all 32 students who completed the Integrated Competence Seminar in Spring, 1980, for whom complete videotaped performances were available.

Instrument

The instrument used in this study of Social Interaction student performance was drawn from the Integrated Competence Seminar (ICS). One part

of this Seminar, Group Discussion, was used to assess student performance on competence levels 1, 2, 3 and 4 of Social Interaction.

Group Discussion was given as a pretest to uninstructed students in order to provide a pre-/posttest comparison for levels 1 to 4 on the Social Interaction Generic Criteria. Thus, uninstructed students experienced a special administration of the same instrument administered to the instructed students validated at level 4 of Social Interaction. While the Social Interaction Generic Criteria were not specifically designed for the ICS Group Discussion, this was the only naturally occurring[1] opportunity to conduct a pre-/posttest study using a similar instrument. Further, some instructed students attempted a level 4 validation. That is, they were contracted for credentialing (contract group). Others completed Group Discussion as part of the ICS requirement, but were not contracted for credentialing at Social Interaction level 4 (no contract group).

The Integrated Competence Seminar (ICS) is an assessment technique usually administered at the mid-point of each Alverno student's college career. The purpose of the ICS is to provide an opportunity for the student to demonstrate how well she can integrate abilities she has developed thus far, and to what extent she can transfer these abilities to a "realistic" situation. Thus, ICS is an out-of-class, cross-disciplinary exercise assessing a student's ability to integrate level 4 of Communications, Problem Solving, Analysis, Social Interaction and Valuing.

As a participant in the ICS assessment, the student takes on the role of a citizen decision-making board member, who must make a selection

---

[1]Studies of this nature are designed to respect curriculum design.

The Integrated Competence Seminar (ICS) is an assessment technique usually administered at the mid-point of each Alverno student's college career. The purpose of the ICS is to provide an opportunity for the student to demonstrate how well she can integrate abilities she has developed thus far, and to what extent she can transfer these abilities to a "realistic" situation. Thus, ICS is an out-of-class, cross-disciplinary exercise assessing a student's ability to integrate level 4 of Communications, Problem Solving, Analysis, Social Interaction and Valuing.

As a participant in the ICS assessment, the student takes on the role of a citizen decision-making board member, who must make a selection among three proposals for funding various city improvements. The student represents one of six areas needing improvement: health, culture, recreation, education, vocational training, or environmental affairs.

The student demonstrates her abilities in a four-hour exercise consisting of three parts: Oral Presentation, In-Basket and Group Discussion. For Oral Presentation, the student studies a set of background information and delivers a speech to persuade the decision-making board to accept her choice among three proposals. She then engages in an In-Basket exercise which derives its name from the in-basket or tray on the manager's desk (containing letters, memos, minutes and reports awaiting action). She solves problems, sets priorities, analyzes, organizes, and makes decisions as if she were encountering a real situation in her office. In Group Discussion, the student meets with other participants to discuss and make decisions on the final proposal, a common plan for the expenditure of funds.

Procedure

The present study examined student performance on one part of the Integrated Competence Seminar (ICS), Group Discussion. This exercise is designed to measure interpersonal and task-oriented skills in Social Interaction. Before the group discussion, students respond in writing to questions about their goals and expectations for accomplishment in the exercise. Each group of five or six students then discusses three funding proposals and attempts to reach consensus on the one which would better serve the needs of the community. The group is initially leaderless. In the course of the half hour discussion, however, leaders emerge. After the exercise is completed, students are asked to reflect on their experience in relation to the other group members' performance.

Following Group Discussion, assessors judge student performance against Social Interaction level 4 criteria. There were some differences in the assessment procedure between the instructed and uninstructed student groups in the present study. The instructed students' (Weekday College) Group Discussions were videotaped and subsequently observed by two assessors who evaluated each student's performance and generated a summary statement describing the student's strengths and weaknesses as an effective group member. This procedure is standardized, and used for assessing Social Interaction level 4 on the ICS.

The uninstructed students' (Weekend College) Group Discussions were audiotaped. During each live group discussion, four to five assessors observed each group interaction. Each assessor observed one student primarily and one student secondarily. The assessors recorded student behaviors during the ongoing discussion. After the discussion, the

likely to be unfamiliar with each other, with little or no prior experience with one another in assessment situations. Students demonstrating their ability in class are likely to interact with students they know, and can more likely predict behavioral responses and patterns in a group situation. Finally, in an assessment situation in class, students are not asked to demonstrate Social Interaction skills in a context that includes all the activities demanded in the Integrated Competence Seminar.

While it is true that uninstructed students performing social interaction skills in the Group Discussion were being credentialed on level 1 of Social Interaction, level 1 validation for incoming students is considered more of an exercise to identify strengths and weaknesses rather than an "assessment." Thus, uninstructed students' level of motivation is assumed to be similar to that of instructed students not contracted for credentialing at level 4.

Social Interaction Generic Criteria

Performance on Social Interaction levels 1, 2, 3 and 4 is coded on four separate dimensions. The Preparation dimension includes criteria for performance prior to Group Discussion: how she perceives her goals, the purpose of the task, etc. (See Table 1). The Preparation dimension criteria are directly related to the student's understanding of the theoretical model underlying behaviors exhibited during Group Discussion (8 criteria).

The Demonstration dimension encompasses all criteria relative to the student's observable social interaction during Group Discussion. Here the student is expected to demonstrate interpersonal as well as task-oriented skills (19 criteria).

insure that uninstructed students would be able to complete Group Discussion.

The Integrated Competence Seminar offers Social Interaction level 4 validations for those students who have not completed all their validations during their regular course work. Thus, some discussion groups were comprised of contracted students who were interacting in a group where each person was being credentialed. These students would seem to operate more out of self-interest and to be less group-oriented. Further, this contracted group may have been comprised of less able students, since they had not completed all Social Interaction validations in class, and had to use the ICS Group Discussion opportunity for their last validation attempt. Consequently, the instructed contracted group may differ from the instructed group not contracted to be credentialed. There may be dissimilar motivational patterns operating which may impact students' performance. The contracted group is assumed to have individual interest invested in their performance in addition to the motivation elicited by the group task. Since the contracted group was performing under the pressure of being credentialed, they may have been less oriented toward group rather than individual discussion outcomes.

The ICS was designed to explore the transfer of abilities learned in class to an out-of-class exercise resembling a more realistic situation. It should be noted however, that transfer of learning from in-class Social Interaction demonstration may differ from the out-of-class ICS demonstration because students are asked not only to demonstrate Social Interaction, but also to integrate these skills with Problem Solving, Analysis and Valuing. Further, the ICS situation is associated with time pressure. Students who perform Social Interaction skills are

assessors evaluated students' performance using evaluation forms for level 1 and level 4 of Social Interaction. The uninstructed students had just entered Alverno College, and were assessed by persons who were previously trained to evaluate student performance of entering students on level 1 criteria only. Therefore, these assessors experienced a brief training session prior to their participation on how to assess for level 4 criteria as well. The assessors evaluating uninstructed students' level 4 performance were probably not as experienced as the level 4 assessors who evaluated instructed students' performance.

The Assessment Center administered the complete Integrated Competence Seminar to the instructed students. Uninstructed students were administered only Group Discussion. Assessment Center personnel are experienced in administering the instruments. However, the administration to the uninstructed students was somewhat different from the instructed students in that the former group was larger than usual. It was also the first time that the Center administered the ICS with so many students at once (n = 90). Usually the Center administers the instrument to 30 students at a time.

All incoming students usually attend a three-week preparation course (IN 050) before attempting competence validation of Social Interaction level 1. In order to control the amount of instruction given the uninstructed group (who were being asked to perform prior to completing course work, during their second day in college), these students participated in a brief introductory session prior to administration of Group Discussion. During this introductory session, the Social Interaction level 1 behavioral categories were introduced and explained (initiating behavior, mediating behavior, challenging behavior, etc.). This was to

The Self-Assessment dimension includes criteria for the student's ability to reflect back on her own experience, to draw conclusions, and to recognize her social interaction strengths and weaknesses (5 criteria).

The Leadership dimension assesses the extent to which the student demonstrates leadership behaviors (0 criteria).

The Social Interaction Generic Criteria are described on an evaluation form created to allow the assessor to score all students similarly for the purpose of the current study (See Table 1).

Procedure for Coding Student Performance

The coder was a faculty member in the Behavioral Sciences Division who was also a member of the Social Interaction Competence Division. The assessor used the Social Interaction Generic Criteria Evaluation Form (Table 1). Because of differences in the data base between uninstructed and instructed students, the coder was aware of which students were instructed and uninstructed. The coder proceeded to score uninstructed group performance by analyzing audiotapes and evaluation forms submitted by assessors (reliability statistics are unavailable). To score instructed group performance, the coder analyzed the videotapes of Group Discussion performance and assessors' summary statements. Thus, the coder did not directly observe uninstructed group performance and occasionally inferred level 4 behaviors from criteria checked by assessors on their separate evaluation forms.

Using the Social Interaction Generic Criteria Evaluation Form, the coder checked all criteria the student demonstrated for the Preparation, Demonstration, and Self-Assessment dimensions. In addition, she rated each student on each dimension on a numerical scale ranging from "not at all effective" (1) to "very effective" (5). The coder did not base her

Table 1
Social Interaction Generic Criteria Evaluation Form

Date _____ Name _____

WDC____
WEC____

PREPARATION (P)

Plans interaction strategy, shows understanding of task goal, shows preparation for content focus of the situation

P 1 2 3 4 5
Not at all effective — Very effective

P_1) ____ perceives the task goal as a problem-solving situation

P_2) ____ identifies own position in relation to the task goal

P_3) ____ describes the key dimensions of the interaction setting

P_4) ____ develops a chronological strategy for assisting the group toward goal achievement

P_5) ____ clearly states the purpose of the task

P_6) ____ relates theory to practice

P_7) ____ cites information from outside sources

P_8) ____ synthesizes the thinking of others

Comments:

DEMONSTRATION (D)

Demonstrates commitment to goal achievement, builds group cohesion, challenges and accepts challenges constructively

D 1 2 3 4 5
Not at all effective — Very effective

D_1) ____ clearly states the purpose of the task

D_2) ____ reviews the task goal whenever appropriate

D_3) ____ summarizes or evaluates group progress in relation to goal achievement

D_4) ____ stimulates discussion through effective goal-related questioning

D_5) ____ shows enthusiasm for goal attainment through non-verbal behavior (facial expression, body posture, voice tone)

D_6) ____ assigns tasks to help implement decisions or attain consensus

D_7) ____ maintains appropriate eye contact with everyone in the group

D_8) ____ uses both person-oriented and task-oriented skills

D_9) ____ uses a broad spectrum of interaction skills

D_{10}) ____ chooses interaction skills appropriate to the situation

D_{11}) ____ attends to non-verbal messages and cues

D_{12}) ____ encourages those who exhibit withdrawing behavior

D_{13}) ____ challenges those who exhibit interfering behavior

D_{14}) ____ mediates differences by searching for common elements

D_{15}) ____ listens to others without interrupting when challenged with opposing points of view

D_{16}) ____ maintains an attitude of openness and responds flexibly when confronted with conflicting ideas

D_{17}) ____ shows ability to modify ideas or opinions when appropriate

D_{18}) ____ contributes a fair share of ideas and suggestions

D_{19}) ____ points out implications of ideas presented

Comments:

SELF-ASSESSMENT (S)

Demonstrates ability to accurately evaluate one's own social interaction effectiveness

S 1 2 3 4 5
Not at all effective — Very effective

S_1) ____ shows how the demonstrated behavior had the intended effect on other(s) as specified in statement of goals -OR-

S_2) ____ provides an adequate rationale for any changes made in the plan

S_3) ____ cites examples of effective behaviors and states how they affected other(s) in the group

S_4) ____ identifies strengths and weaknesses of own performance

S_5) ____ points out ways to improve one's own effectiveness

Comments:

LEADERSHIP (L)[a]

L 1 2 3 4 5
Not at all effective — Very effective

Comments:

[a] e.g. P_6, D_1, S_4

judgment on the number of criteria already checked for each dimension, but rather on her overall impression of the student's performance. The Leadership dimension was scored on a similar numerical scale, based on the coder's subjective judgment of which behaviors demonstrate leadership. A student was identified as mastering a dimension if she was rated either a 3, 4 or 5 on the numerical scale.

Design

The study design is to compare uninstructed students' performance to instructed students' performance on Social Interaction Generic Criteria assessing levels 1, 2, 3 and 4 of the Social Interaction competence, a two group comparison.

Subsequent comparisons within each of the two groups were also planned. Following data collection, it became clear that uninstructed students entering Weekend College varied on age and prior college experience. Consequently, students were categorized as "older" if they were age 27 to 54 years, and "younger" if they were age 19 to 25 years. The uninstructed group differed in the amount of prior college experience, and were categorized as having prior college experience (range, 30 to 120 credits) or no prior college (0 to 24 credits). Thus, three groups were identified: Group I (WEC, uninstructed, older, prior college); Group II (WEC, uninstructed, older, no prior college); and Group III (WEC, uninstructed, younger, no prior college). In order to control for age, the no prior college category included 9 students who had some prior college (18 to 24 credits). Since the number of credits for these younger students was at the lower end of the distribution of credits, the cutoff point for no prior college was set at 24 or less credits. The Weekday College instructed group was also examined for variation in

age (18 to 49 years) and prior college experience (0 to 73 credits). Table 2 shows the number of students in each group, the mean, median and range for age, and the mean, median and range for number of prior credits.

Examination of Table 2 indicates that Group I (WEC, uninstructed, older, prior college) and Group II (WEC, uninstructed, older, no prior college) are comparable with respect to median age, but are dissimilar on median college experience. Consequently, a comparison of Group I and II can examine the extent to which prior college experience affects uninstructed students' performance on the Social Interaction Generic Criteria.

Further, Group II (WEC, uninstructed, older, no prior college) and Group III (WEC, uninstructed, younger, no prior college) students are somewhat comparable in that Group II uninstructed students have no prior college experience, while Group III has an average of 8 credits with a median of 18 credits. Group II and III are dissimilar in age with 20 years difference in the mean age and 10 years difference in median age. This suggests that a comparison between Group II and Group III could control somewhat for prior college experience, and allow us to examine the extent to which maturation (age) affects uninstructed students' performance on the Social Interaction Generic Criteria.

The Weekday College instructed group provides us an opportunity to examine the difference in performance between students contracted for credentialing at level 4 of Social Interaction ($\underline{n}$ = 9), and those who were not contracted ($\underline{n}$ = 23). Examination of Table 2 shows that median age and median prior college experience are almost identical for Group III (WEC, uninstructed, younger, no prior college) uninstructed students and for Group IV (WDC, instructed, younger, no prior college, no contract) instructed students. This group comparison may provide

Table 2

Mean, Median and Range for Age and Prior College Experience (credits) for Uninstructed and Instructed Students Categorized for Planned Comparisons

Group	n	Age			Prior College Credits		
		Range	M	Mdn	Range	M	Mdn
Group I Weekend College (WEC) Uninstructed Older Prior college	27	27-44	38.5	33	30-120	63.0	42
Group II Weekend College (WEC) Uninstructed Older No prior college	25	27-54	42.9	33	0	0	0
Group III Weekend College (WEC) Uninstructed Younger No prior college	18	19-25	22.5	23	0-24	8.2	18
Group IV Weekday College (WDC) Instructed Younger No prior college No contract	23	18-49	27.5	24	0-50	20.9	18
Group V Weekday College (WDC) Instructed Younger No prior college Contract	9	19-40	30.2	27	0-73	26.8	23

the best test of the effects of instruction on social interaction performance. Further, Groups II and III compared with Groups IV and V may provide a better comparison of the effects of instruction, since Group I has had extensive prior college experience.

Following examination of the effects of instruction, age and prior college experience, group performance will be examined to investigate the extent to which patterns in performance can be identified. These analyses will be used to interpert the meaning of the Social Interaction competence as a construct. While the groups may be somewhat different on several variables (age, prior college, instruction), students may behave similarly in response to the Social Interaction criteria. Different patterns in performance may suggest combinations of criteria that can illuminate the nature and meaning of the competence.

Analysis

To examine the differences and interactions between the groups, t-tests and ANOVA were employed. Discriminant analysis and Pearson product moment correlation matrices were used to investigate patterns of behaviors. Multiple regression and intra-item consistency tables examined the contribution of the behavioral criteria to the several dimensions of the Social Interaction competence.

RESULTS

The data analysis plan for the Social Interaction competence study was presented to the Social Interaction Division by the Office of Research Evaluation. The guidelines and questions for the data analysis were discussed and a set of priorities for the analysis were generated.

The Division decided to explore the following issues as a first priority:

- How do instructed students perform compared to uninstructed students on each of the four Social Interaction dimensions (instructional validity)?
- How do instructed students perform compared to uninstructed students on each of the behavioral criteria that comprise the dimensions (instrument criteria validity)?
- How might we better understand the meaning of the Social Interaction competence from an analysis of students' Social Interaction performance, irrespective of differences between instructed and uninstructed students (construct validity)?

Instructional Validity

The purpose of the initial data analysis is to examine the effects of instruction on student performance at Social Interaction level 4. We examined differences in performance between instructed and uninstructed student groups by comparing and contrasting Weekday and Weekend College students who differ in instructional experience, age, and prior college experience on each one of the Social Interaction dimensions (Preparation,

Demonstration, Self-Assessment and Leadership). We then identified performance within Weekday College instructed groups which may be attributed to instruction.

Figures 1 through 5 describe the relative frequency distributions of dimension scores for each of the student groups compared. As mentioned earlier, a dimension score is comprised of a 5-point scale which ranges from 1 (not at all effective) to 5 (very effective). A mastery cut-off score at point 3 of the scale was chosen by the Social Interaction faculty scorer. Students who scored 3, 4 or 5 were classified by the scorer as mastery students and students who scored 1 or 2 were classified as no mastery students.

Table 3 presents the percentages of mastery and no mastery students per group by dimension (Preparation, Demonstration, Self-Assessment and Leadership). Statistical significance tests of the difference in proportions (Glass and Stanley, 1970) were employed to compare each group with each other group on mastery performance for three of the dimension scores (Leadership was not included).[1] The only significant differences in proportions of students who mastered each dimension were found in the Preparation dimension. No significant differences were found in the Demonstration or Self-Assessment dimensions. In the Preparation dimension, Group I (WEC, uninstructed, older, prior college) had a significantly higher percentage of mastery students ($\underline{P}$ = 59.26) compared to Group II (WEC, uninstructed, older, no prior college) ($\underline{P}$ = 32.00)

---

[1]Since proportions were tested, results from mastery students will be identical to results from no mastery students; thus, the latter were not compared.

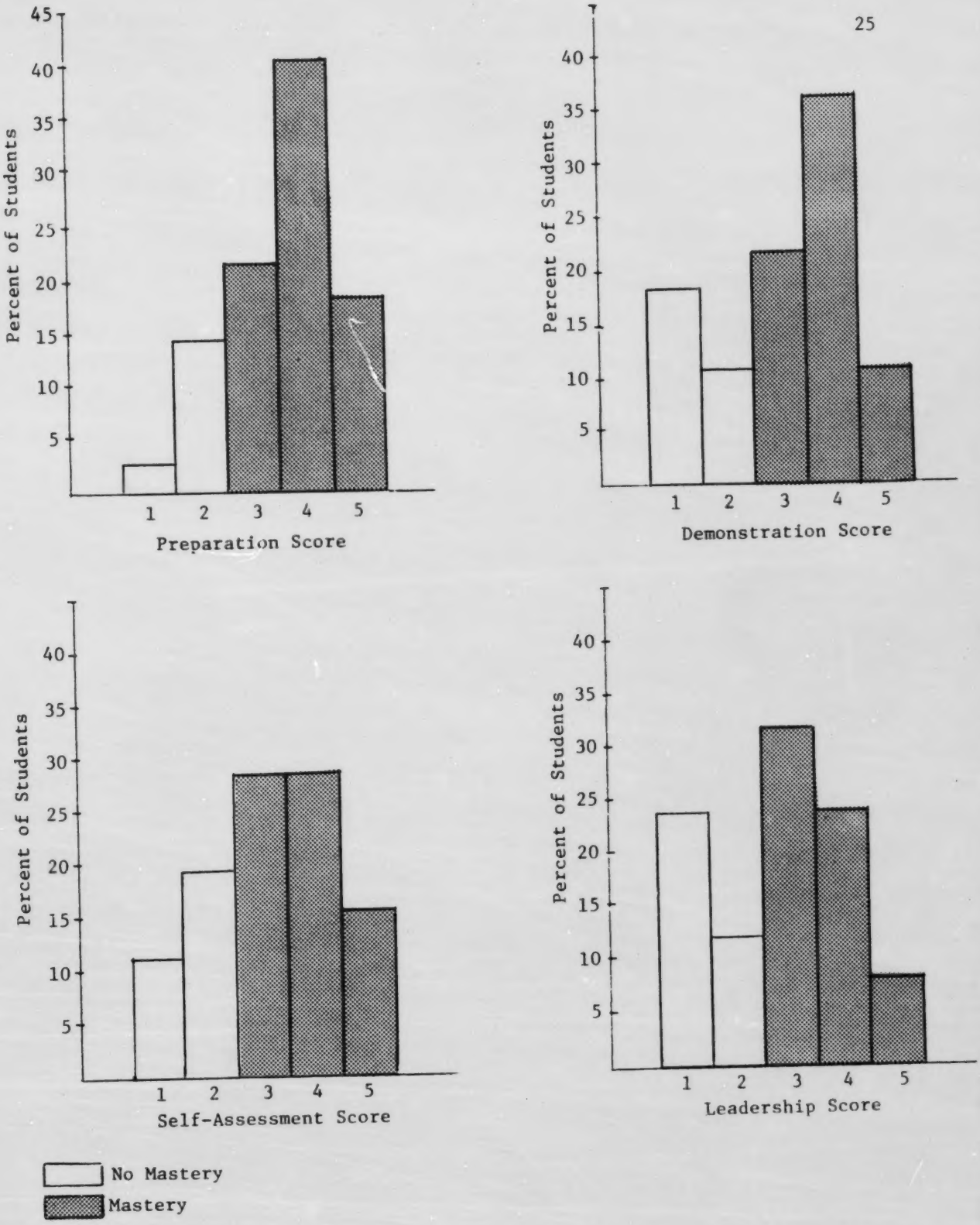

Figure 1. Distributions of Integrated Competence Seminar scores for Group I (WEC, uninstructed, older, prior college): Mastery and No Mastery students.

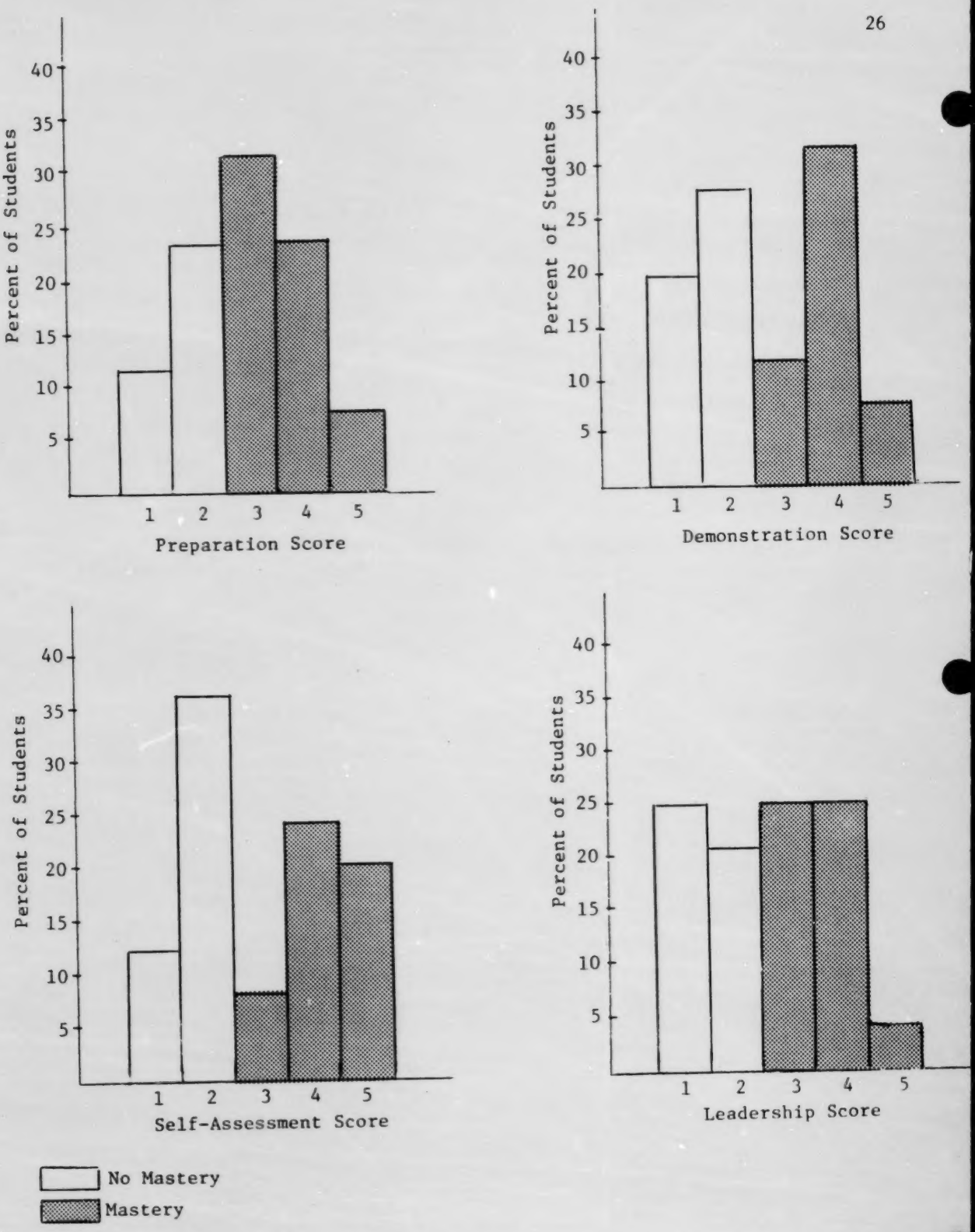

Figure 2. Distributions of Integrated Competence Seminar scores for Group II (WEC, uninstructed, older, no prior college): Mastery and No Mastery students

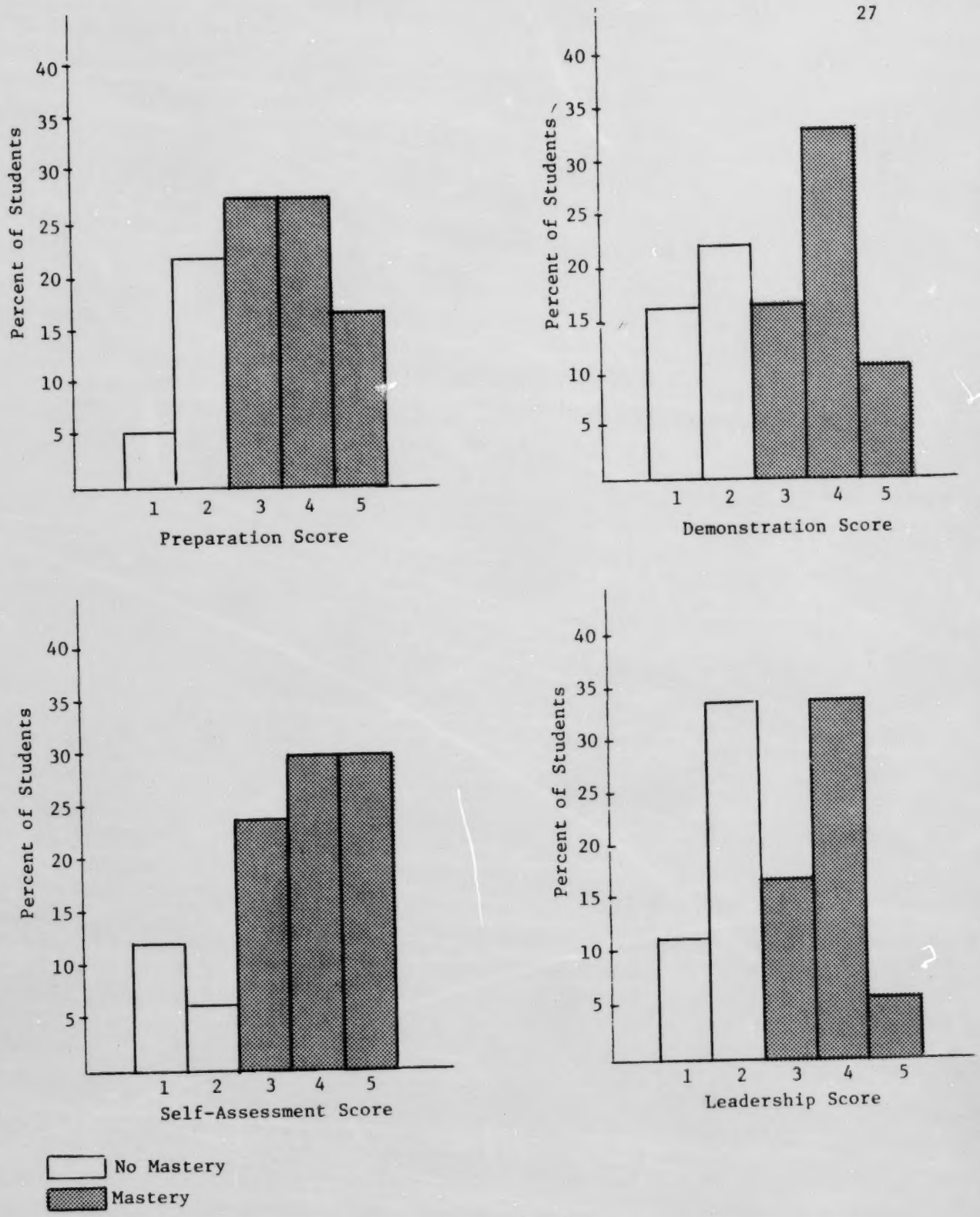

Figure 3. Distributions of Integrated Competence Seminar scores for Group III (WEC, uninstructed, younger, no prior college): Mastery and No Mastery students.

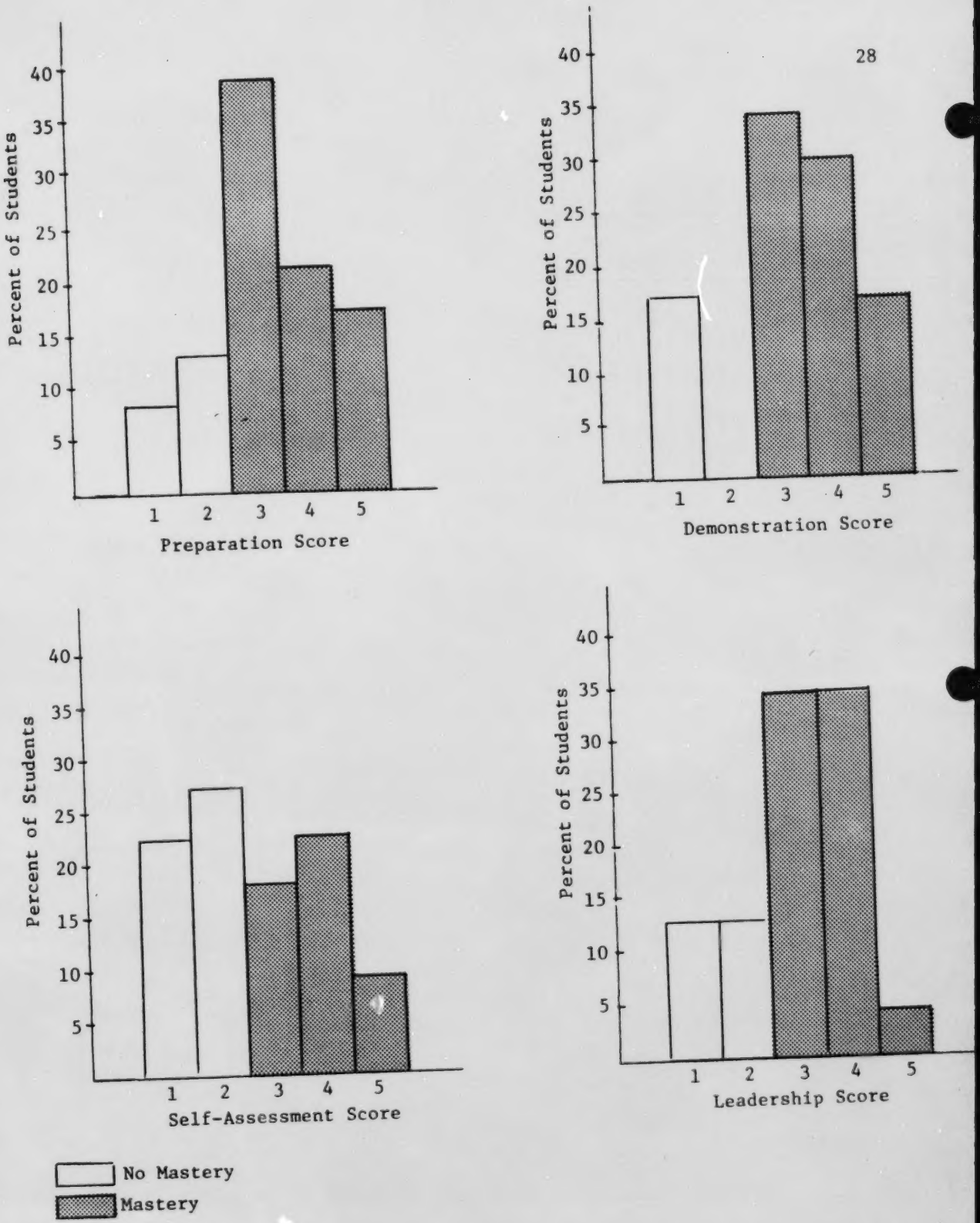

Figure 4. Distributions of Integrated Competence Seminar scores for Group IV (WDC, instructed, younger, no prior college, no contract): Mastery and No Mastery students.

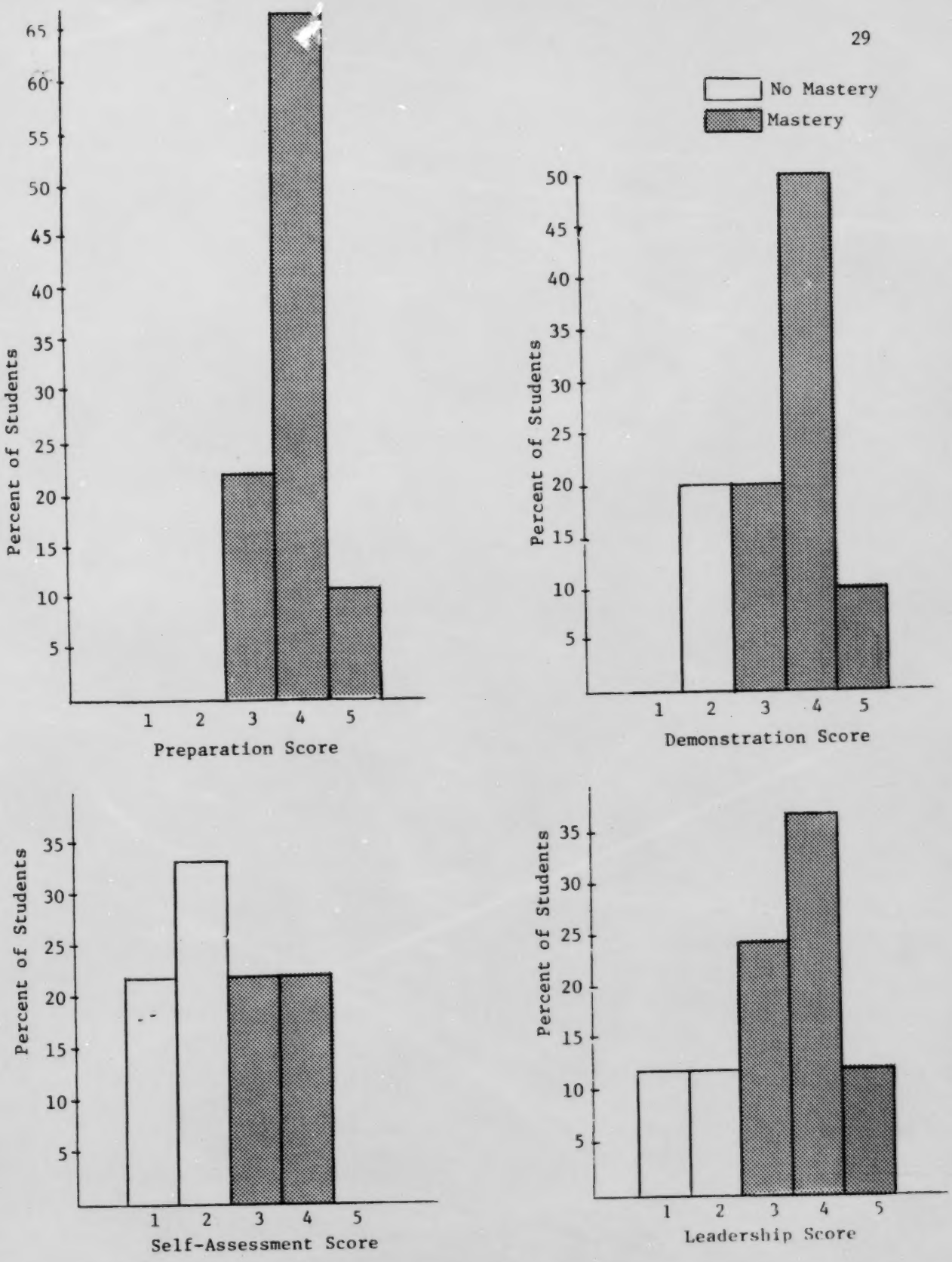

Figure 5. Distributions of Integrated Competence Seminar scores for Group V (WDC, instructed, younger, no prior college, contract): Mastery and No Mastery students.

Table 3

Percentage of Mastery and No Mastery Students per Group by Social Interaction Dimension

Students	Group I WEC Uninstructed Older Prior college	Group II WEC Uninstructed Older- No prior college	Group III WEC Uninstructed Younger No prior college	Group IV WDC Instructed Younger No prior college No contract	Group V WDC Instructed Younger No prior college Contract
			Preparation		
No Mastery	40.74	68.00	55.56	60.87	22.22
Mastery	59.26	32.00	44.44	39.13	77.78
			Demonstration		
No Mastery	51.85	60.00	55.56	52.17	40.00
Mastery	48.15	40.00	44.44	47.83	60.00
			Self-Assessment		
No Mastery	57.69	56.00	41.18	68.18	77.78
Mastery	42.31	44.00	58.82	31.82	22.22
			Leadership		
No Mastery	68.00	70.83	61.11	60.87	50.00
Mastery	32.00	29.17	38.89	39.13	50.00

($z = 1.97$, $p < .05$). The Weekend College groups are uninstructed whereas the Weekday groups are instructed students. The Group III (WEC, uninstructed, younger, no prior college), and Group IV (WDC, instructed, younger, no prior college, no contract) comparison was examined more carefully, since Group III and IV seem to be the most comparable of all groups. The effectiveness of instruction can, of course, be more confidently explored when the groups compared are similar. Groups III (WEC, uninstructed, younger, no prior college), and Group IV (WDC, instructed, younger, no prior college, no contract) (which allow for a better comparison of instructed and uninstructed students), did not differ significantly with respect to mastery in the Preparation dimension.

Group V (WDC, instructed, younger, no prior college, contract) had a significantly higher percentage of mastery students ($P = 77.78$) compared to the other instructed students in Group IV (WDC, instructed, younger, no prior college, no contract) ($P = 39.13$) ($z = 1.97$, $p < .05$), and Group II (WEC, uninstructed, older, no prior college) ($P = 32.00$) ($z = 2.37$, $p < .05$). Thus, Weekday students who contracted for credentialing (Group V: WDC, instructed, younger, no prior college, contract) demonstrated more effective Preparation compared to the Weekday students who were not contracted for credentialing (Group IV: WDC, instructed, younger, no prior college, no contract).

Following comparison of students who mastered each of the dimensions, we examined the range of dimension scores (1-5) across all the groups. Table 4 shows the means and standard deviations per group by dimension. Paired t-tests were employed to investigate the mean differences between pairs of groups on each of the dimension scores. Table 5 shows the level of significance of t-test comparisons between groups.

Table 4

Group Means and Standard Deviations for Each Social Interaction Dimension

	Group I WEC Uninstructed Older Prior college	Group II WEC Uninstructed Older No prior college	Group III WEC Uninstructed Younger No prior college	Group IV WDC Instructed Younger No prior college No contract	Group V WDC Instructed Younger No prior college Contract
Students	n = 27	n = 25	n = 18	n = 23	n = 9
			Preparation		
Mean	3.56	2.92	3.28	3.26	3.89
SD	1.09	1.15	1.19	1.18	0.60
			Demonstration		
Mean	3.11	2.80	3.00	3.30	3.50
SD	1.31	1.32	1.32	1.29	0.97
			Self-Assessment		
Mean	3.15	3.04	3.59	2.68	2.44
SD	1.26	1.40	1.33	1.32	1.13
			Leadership		
Mean	2.80	2.63	2.89	3.04	3.25
SD	1.29	1.24	1.18	1.11	1.28

Table 5

Group Comparisons Showing Significant Differences from t-Tests for Each Dimension

Group	DIMENSION														
	Preparation					Demonstration					Self-Assessment				
	I	II	III	IV	V	I	II	III	IV	V	I	II	III	IV	V
Group I WEC Uninstructed Older Prior college															
Group II WEC Uninstructed Older No prior college	2.04*					.85					.31				
Group III WEC Uninstructed Younger No prior college	.80	-.99				.28	-.49				-1.07	-1.29			
Group IV WDC Instructed Younger No prior college No contract	.91	-1.01	.05			-.52	-1.33	-.74			1.27	.90	2.12*		
Group V WDC Instructed Younger No prior college Contract	-1.15	2.39*	-1.78	-1.98*		-.98	-1.73	-1.14	-.48		1.58	1.27	2.20*	.50	

Note. *$p < .05$

As Table 5 indicates, significant differences between groups on mean dimension scores for Preparation were not different than the results obtained from tests of proportional differences. Group V (WDC, instructed, younger, no prior college, contract), was significantly more effective in Preparation ($\underline{M}$ = 3.89) compared to Group II (WEC, uninstructed, older, no prior college) ($\underline{M}$ = 2.92), ($\underline{t}(41) = 2.39$, $\underline{p} < .05$) and Group IV (WDC, instructed, younger, no prior college, no contract) ($\underline{M}$ = 3.26; $\underline{t}(39) = -1.98$, $\underline{p} < .05$). In the previous comparison of mastery students Weekend College, uninstructed, older women with extensive prior college experience (Group I) are similar in Preparation effectiveness to Group V (WDC, instructed, younger, no prior college, contract). The WEC uninstructed group (Group II: WEC, uninstructed, older, no prior college) ($\underline{M}$ = 2.92), without prior college experience, is significantly less effective ($t(41) = 2.02$, $\underline{p} < .05$) than the WEC uninstructed Group I (WEC, uninstructed, older, prior college) ($\underline{M}$ = 3.56).

There were significant mean differences on the Self-Assessment dimension which did not appear in the tasks of proportional difference in mastery. Weekend young students (Group III: WEC, uninstructed, younger, no prior college) were significantly more effective: Self-Assessment of their performance ($\underline{M}$ = 3.59) compared to any of the instructed groups (Group IV, $\underline{M}$ = 2.68) (Group V, $\underline{M}$ = 2.44) ($\underline{t}$ (39) = 2.12 $\underline{p} < .05$; $\underline{t}(34) = 2.20$ $\underline{p} < .05$, respectively).

Since Group III (WEC, uninstructed, younger, no prior college) is somewhat similar to Group IV (WDC, instructed, younger, no prior college, no contract), we may conclude, with respect to Self-Assessment performance, that instructional validity has not been demonstrated via most of these comparisons.

One-way repeated measures ANOVAS were employed to investigate the main effect of group by dimension as well as the interaction effect between group and dimension. Such analyses assist our understanding of how group differences in Social Interaction performance are affected by the variables (instructed/uninstructed; older/younger; contract/no contract).

Repeated measures ANOVAS (2 levels of group by four levels of dimension)[1] yielded no group main effects among all five paired comparisons. The group comparison which best examines the instructed/uninstructed comparison, Group III (WEC, uninstructed, younger, no prior college) and Group IV (WDC, instructed, younger, no prior college, no contract), resulted in an interaction effect (F (17,22) = 4.75, $p < .01$) which indicates that the direction of the differences between the two groups on each of the four dimensions is not consistent across all four dimensions. This interaction effect (see Figure 6) may indicate the existence of other factors which affect student performance other than group membership and the dimension being performed. When dimension means for Group III (WEC, uninstructed, younger, no prior college) and Group IV (WDC, instructed, younger, no prior college, no contract) are examined, it is clear that the direction of group performance means varies inconsistently with type of dimension. Figure 6 shows that Group IV (WDC, instructed, younger, no prior college, no contract) performed significantly lower than Group III (WEC, uninstructed, younger, no prior

---

[1] A two-group comparison was used rather than a 5 x 4 design to investigate possible interaction effects immediately, but also because the groups compared differ from each other with respect to the construct being investigated (instructed/uninstructed;older/younger; contract/no contract).

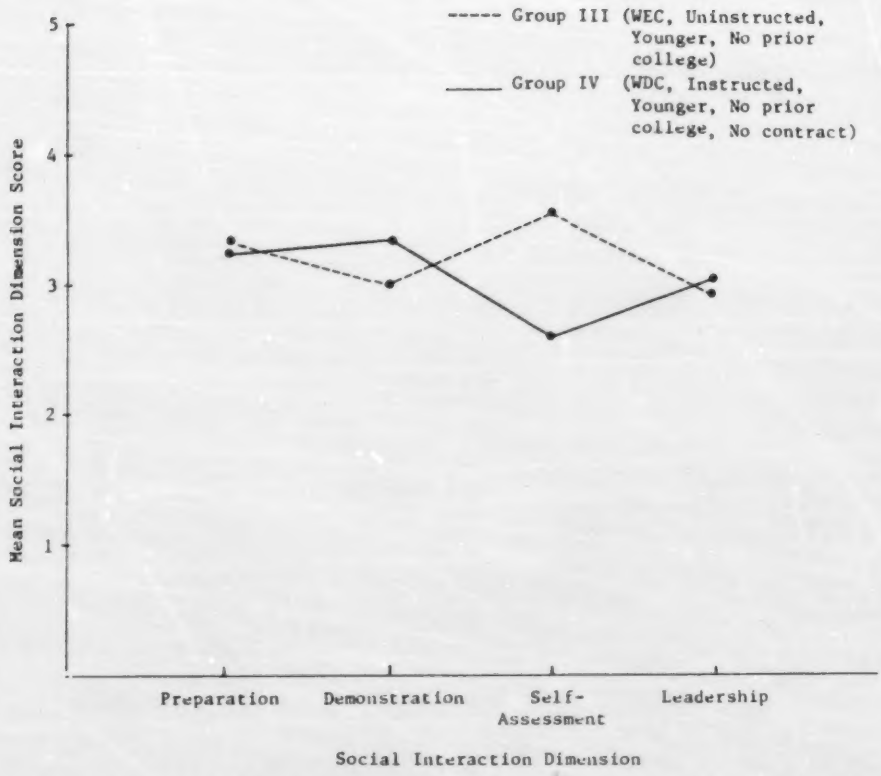

Figure 6. Interaction Effect Between Group III and Group IV

college) on Self-Assessment but Group IV is higher (although nonsignificant) than Group III on Demonstration and Leadership.

In the other paired comparisons, a significant dimension effect (F (22,8) = 8.78, $p < .001$) was obtained for Groups IV (WDC, instructed, younger, no prior college, no contract) and V (WDC, instructed, younger, no prior college, contract) and Groups I (WEC, uninstructed, older, prior college) and IV (WDC, instructed, younger, no prior college, no contract) (F (26, 22) = 3.77, $p < .05$).

In sum, there were no group main effects among the five comparisons. Students do not perform differently given their group membership. But there were two significant dimension main effects in two of the five comparisons. What differences in performance exist seem to be explained by the dimension the student is performing. Dimension affects student performance more than group membership.

When dimension means were collapsed across Group IV (WDC, instructed, younger, no prior college, no contract) and Group V (WDC, instructed, younger, no prior college, contract), Preparation (M = 3.45) was equal to Demonstration (M = 3.42). Both were significantly higher than Self-Assessment (M = 2.62). When dimension means were collapsed across Group I (WEC, uninstructed, older, prior college) and Group IV (WDC, instructed, younger, no prior college, no contract), Preparation (M = 3.44) was not significantly higher than Demonstration (M = 3.25). Both were significantly higher than Self-Assessment (M = 2.94).

Instrument Criteria Validity

The second phase of the data analysis focused on the evaluation of the instrument criteria.

- To what extent do the instrument criteria measure the effects of instruction?
- Which criteria best discriminate instructed from uninstructed groups?
- Are all important Social Interaction skills at level 4 represented by the instrument's behavioral criteria?
- Which criteria contribute most to effective Preparation, Demonstration, or Self-Assessment?

In order to identify criteria that discriminate between instructed and uninstructed groups, we compared the frequency of response per criterion in one group with that of another via tests of association (chi-square). Table 6 shows the percentage of students who responded to each criterion within each group. Eight two-group comparisons, and the significant chi-squares per criteria comparison are also indicated. We found that few criteria were significantly associated within each two-group comparison.

In the first comparison, Group III (WEC, uninstructed, younger, no prior college) vs. Group IV (WDC, instructed, younger, no prior college, no contract), the uninstructed group performed significantly higher on the following criteria:

P4: Develops a chronological strategy for assisting the group toward goal achievement

D14: Mediates differences by searching for common elements

S4: Identifies strengths and weaknesses of own performance

S5: Points out ways to improve one's own effectiveness

Table 6

Criterion Response Frequency Within Two-Group Comparisons

(continued on next page)

Group Comparison	Criterion																	
	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}
III	94.4	94.4	00.0	50.0	83.3	11.1	83.3	11.1	50.0	27.8	44.4	33.3	66.7	22.2	94.4	16.7	16.7	27.8
IV	91.3	73.9	00.0	8.7	73.9	21.7	95.7	17.4	43.5	26.1	30.4	69.6	73.9	39.1	95.7	00.0	13.0	21.7
$\underline{x}^2$				6.80**								3.97*						
I	88.9	88.9	00.0	48.1	77.8	25.9	88.9	40.7	66.7	25.9	18.5	77.8	81.5	29.6	74.1	14.8	29.6	33.3
IV	91.3	73.9	00.0	8.7	73.9	21.7	95.7	17.4	43.5	26.1	30.4	69.6	73.9	39.1	95.7	00.0	13.0	21.7
$\underline{x}^2$				7.42**														
II	84.0	88.0	00.0	40.0	68.0	8.0	80.0	24.0	56.0	20.0	36.0	64.0	60.0	24.0	60.0	12.0	20.0	36.0
IV	91.3	73.9	00.0	8.7	73.9	21.7	95.7	17.4	43.5	26.1	30.4	69.6	73.9	39.1	95.7	00.0	13.0	21.7
$\underline{x}^2$				4.70 *											6.72**			
IV	91.3	73.9	00.0	8.7	73.9	21.7	95.7	17.4	43.5	26.1	30.4	69.6	73.9	39.1	95.7	00.0	13.0	21.7
V	90.0	90.0	10.0	10.0	80.0	20.0	90.0	60.0	50.0	30.0	40.0	80.0	100.0	60.0	100.0	10.0	10.0	30.0
$\underline{x}^2$								4.14 *										

*$\underline{p} < .05$
**$\underline{p} < .01$

Table 6 Continued

Group Comparison	Criterion													
	D_{11}	D_{12}	D_{13}	D_{14}	D_{15}	D_{16}	D_{17}	D_{18}	D_{19}	S_1	S_2	S_3	S_4	S_5
III	77.8	16.7	22.2	27.8*	55.6	77.8	61.1	83.3	55.6	38.9	44.4	72.2	83.3	38.9
IV	91.3	13.0	21.7	00.0	60.9	56.5	52.2	87.0	73.9	60.9	26.1	60.9	43.5	00.0
$\underline{x}^2$				4.90 *									5.16*	8.21**
I	51.9	18.5	14.8	11.1	55.6	44.4	51.9	85.2	51.9	51.9	33.3	63.0	81.5	33.3
IV	91.3	13.0	21.7	00.0	60.9	56.5	52.2	87.0	73.9	60.9	26.1	60.9	43.5	00.0
$\underline{x}^2$	7.42 **												6.22*	7.22**
II	56.0	24.0	20.0	16.0	48.0	48.0	44.0	84.0	52.0	48.0	24.0	60.0	80.0	32.0
IV	91.3	13.0	21.7	00.0	60.9	56.5	52.2	87.0	73.9	60.9	26.1	60.9	43.5	00.0
$\underline{x}^2$	5.88*												5.35*	6.68**
IV	91.3	13.0	21.7	00.0	60.9	56.5	52.2	87.0	73.9	60.9	26.1	60.9	43.5	00.0
V	80.0	00.0	20.0	60.0	60.0	20.0	70.0	90.0	50.0	70.0	00.0	40.0	20.0	00.0
$\underline{x}^2$				13.07 ***										

* $\underline{p} < .05$
** $\underline{p} < .01$
*** $\underline{p}$ .001

The instructed group performed higher on the following criterion:

D4: Stimulates discussion through effective goal related questions

In this comparison, no discriminative criteria were identified that could be attributed to instruction.

In the second comparison, Group I (WEC, uninstructed, older, prior college) vs. Group IV (WDC, instructed, younger, no prior college, no contract), the uninstructed students performed significantly higher on the following criteria:

P4: Develops a chronological strategy for assisting the group toward goal achievement

S4: Identifies strengths and weaknesses of own performance

S5: Points out ways to improve one's own effectiveness

The instructed students performed higher on the following criterion:

D11: Attends to nonverbal messages and cues

In the third comparison, Group II (WEC, uninstructed, older, no prior college) vs. Group IV (WDC, instructed, younger, no prior college, no contract), the uninstructed students performed higher on the following criteria:

P4: Develops a chronological strategy for assisting the group toward goal achievement

S4: Identifies strengths and weaknesses of own performance

S5: Points out ways to improve one's own effectiveness

Instructed students performed higher on these criteria:

D7: Maintains appropriate eye contact with everyone in the group

D11: Attends to non-verbal messages and cues

In the last comparison, between the two Weekday College Groups (IV and V), the contracted group performed significantly higher on these criteria:

P8: Synthesizes the thinking of others

D14: Mediates differences by searching for common elements

All three uninstructed WEC groups Groups I, II and III performed significantly higher on the following criteria:

P4: Develops a chronological strategy for assisting the group toward goal achievement

S4: Identifies strengths and weaknesses of own performance

S5: Points out ways to improve one's own effectiveness

The instructed, no contract group (Group IV) performed consistently higher on one criterion:

D11: Attends to non-verbal messages and cues

If all Weekend College students are compiled into one uninstructed group, and all Weekday College students into one instructed group (See Table 7), the discriminative criteria that reflect the effects of instruction are:

P4: Develops a chronological strategy for assisting the group toward goal achievement

P7: Cites information from outside sources

D6: Assigns tasks to help implement decisions or attain consensus

D7: Maintains appropriate eye contact with everyone in the group

D11: Attends to non-verbal messages and cues

S4: Identifies strengths and weaknesses of own performance

The dimension scores for effective Preparation, Demonstration and Self-Assessment were entered in a multiple regression analysis to investigate which criterion contributes the most to variation in the dimension

Table 7

All Instructed and Uninstructed Student Performance on Criteria Compared

Criterion	Uninstructed Group (n=70)		Instructed Group (n=32)		z Score
	Number responding	Percent responding	Number responding	Percent responding	
P_1	65	92 %	30	93 %	- .17
P_2	57	81 %	23	71 %	1.09
P_3	0	0 %	1	3 %	-1.49
P_4	32	45 %	3	9 %	3.59 **
P_5	53	75 %	25	78 %	- .27
P_6	11	15 %	7	21 %	- .76
P_7	59	84 %	32	100 %	-2.37 *
P_8	19	27 %	10	31 %	- .43
D_1	31	44 %	15	46 %	- .24
D_2	17	24 %	9	28 %	- .41
D_3	22	31 %	11	34 %	- .30
D_4	43	61 %	24	75 %	-1.34
D_5	49	70 %	27	84 %	-1.55
D_6	18	25 %	15	46 %	-2.12 *
D_7	52	74 %	32	100 %	-3.16 **
D_8	10	14 %	1	3 %	1.69
D_9	16	22 %	4	12 %	1.22
D_{10}	23	32 %	8	25 %	.80
D_{11}	42	60 %	29	90 %	-3.12 **
D_{12}	14	20 %	3	9 %	1.34
D_{13}	13	18 %	7	21 %	- .39
D_{14}	12	17 %	6	18 %	- .20
D_{15}	37	52 %	20	62 %	- .91
D_{16}	38	54 %	15	46 %	.70
D_{17}	36	51 %	19	59 %	- .75
D_{18}	59	84 %	29	90 %	- .86
D_{19}	37	52 %	22	68 %	-1.51
S_1	33	47 %	21	65 %	-1.74
S_2	23	32 %	6	18 %	1.47
S_3	45	64 %	18	56 %	.77
S_4	57	81 %	12	37 %	4.40 **
S_5	24	34 %	0	0 %	3.79

*$p < .05$
**$p < .01$

score. Each criterion was considered as an independent variable and entered in a step-wise regression. The dimension scores were used as the predictor variables. Table 8 shows the multiple correlation R, R^2 and the ranks of the beta weights for each entered criterion.

Table 8

Multiple Regression Summary Table for Behavioral Criteria Predicting the Preparation Dimension Score

Step	Criterion		Multiple R	R^2	Beta in final step	Simple r
1	P_5	Clearly states the purpose of the task	.59	.35	.40	.59
2	P_4	Develops a chronological strategy for assisting the group toward goal achievement	.70	.49	.31	.46
3	P_1	Perceives the task goal as a problem-solving situation	.77	.60	.26	.51
4	P_7	Cites information from outside sources	.81	.66	.24	.40
5	P_6	Relates theory to practice	.84	.71	.24	.42

Since no significant differences were found among each criterion from one group compared to the same criterion in other groups, all students across the five groups were combined in the multiple regression analysis (n = 99).

The five criteria entered in the step-wise regression explain 70% of the variance in the Preparation dimension score. Criterion "clearly states the purpose of the task (P5)" contributes the most with 35% of the variance explained. Criterion "perceives the task goal as a problem solving situation (P1)," accounts for 11% of the variance. Criterion "cites information from outside sources (P7)" accounts for 6%, and criterion "relates theory to practice (P6)" accounts for 5%.

The simple r indicates the correlation between the criterion and the dimension score, and is a measure of the internal consistency of the criterion. Thus, behavioral criteria which correlate highly with the dimension score can also be considered to be discriminative criteria. Criteria "clearly states the purpose of the task (P5)" and "perceives the task goal as a problem solving situation (P1)" are thus good discriminative criteria (P5, r = .59; P1, r = .51).

Table 9

Multiple Regression Summary Table for Behavioral Criteria Predicting the Demonstration Dimension Score

Step	Criterion		Multiple R	R^2	Beta in final step	Simple r
1	D_4	Stimulates discussion through effective goal related questioning	.66	.43	.32	.66
2	D_5	Shows enthusiasm for goal attainment through nonverbal behavior	.75	.56	.30	.53
3	D_{16}	Maintains an attitude of openness and responds flexibly when confronted with conflicting ideas	.80	.64	.26	.45
4	D_3	Summarizes or evaluates progress in relation to goal achievement	.84	.70	.26	.43
5	D_{18}	Contributes a fair share of ideas and suggestions	.87	.75	.26	.58

The five criteria entered in Table 9 explain 75% of the variance in the Demonstration dimension score. Criterion "stimulates discussion through effective goal related questioning (D4)" explains 43% of the Demonstration dimension variance and is also a good discriminative criterion (r = .66). Criterion "contributes a fair share of ideas and suggestions (D18)" and "shows enthusiasm for goal attainment through nonverbal behavior (D5)" are also good discriminative criteria, although D5 explains 12% and D18 only 5% of the variance.

The five criteria entered in Table 10 explain 77% of the variance in the Self-Assessment dimension score. Criterion "cites examples of effective behaviors and states how they affected others in the group (S3)" and "identifies strengths and weaknesses of own performance (S4)" account for 35% and 28% of the variance respectively and they are also good discriminative items (S3, r = .60; S4, r = .59).

Table 10

Multiple Regression Summary Table for Behavioral Criteria Predicting the Self-Assessment Dimension Score

Step	Criterion		Multiple R	R^2	Beta in final step	Simple r
1	S_3	Cites examples of effective behaviors and states how they affected other(s) in the group	.60	.36	.46	.60
2	S_4	Identifies strengths and weaknesses of own performance	.80	.64	.41	.59
3	S_5	Points out ways to improve one's own effectiveness	.85	.72	.35	.47
4	S_1	Shows how the demonstrated behavior had the intended effect on other(s)	.86	.75	.33	.30
5	S_2	Provides an adequate rationale for any changes made in the plan	.88	.77	.23	-.03

Construct Validity

This section of the analysis explores the construct validity of the Social Interaction competence. How best is Social Interaction as an ability defined?

Questions representing some of the construct validity issues are as follows:

- Can we identify patterns of social interaction behaviors which characterize each group of students? Does the Social Interaction construct as a defined set of abilities have the same meaning for different groups? Are different groups characterized by different Social Interaction skills?
- As we review patterns of performance, how can we describe the entry level of Social Interaction competence among uninstructed WEC students who may have had informal learning experiences?
- How do the patterns of performance in students who are uninstructed compare to students who have been instructed?

These questions allow us to look at the competence more holistically, rather than looking at specific dimension scores as we did in the previous univariate analysis.

Multivariate analysis was used to examine these construct validity issues. A discriminant analysis (Klecka, 1975) was used to simultaneously examine all the Social Interaction criteria across the five groups. By exploring similarities and differences among Social Interaction behaviors across five different groups simultaneously, groups which vary with respect to instruction, age and other college experience, we can more broadly describe the range of meaning of the Social Interaction competence.

We are attempting to identify the behaviors that best separate groups from one another. Once these behaviors are identified, we can use them to describe a group. If each group is described by a different set of behaviors, we may infer that each group behaves differently because they conceptualize Social Interaction differently.

Each behavioral criterion (0 or 1) including each dimension score (1, 2, 3, 4 or 5) was included in the analysis. A total of 36 variables (criteria) were entered for each one of the five groups, three Weekend College groups (I, II, and III) and two Weekday College groups (IV and V). Four discriminant functions[1] were generated. Table 11 summarizes the statistics generated for one of the four functions.

Table 11 indicates that functions 1 and 2 account for 82% of the variance and they both contribute significantly to the separation among the groups (function 1, $p < .000$); (function 2, $p < .008$). Thus only functions 1 and 2 are used to describe the groups.

---

[1]A discriminant function is a set of numbers that can have either positive or negative values (coefficients). Each criterion is assigned a value which indicates the extent to which the criterion contributes to the similarities and differences between groups. The higher the value or coefficient, positive or negative, the greater the contribution of the criterion to the separation of the groups from each other. A positive or negative value indicates only the direction of the coefficient; it is not an index of the quality of performance. From a five group comparison, four functions can be generated. Those functions that are statistically significant are selected.

Table 11

Canonical Discriminant Functions for Five Groups Which Differ in Program, Instruction, Age and Other College Experience

Function	Eigenvalue	Percent of Variance	Cumulative Percent of Variance	Canonical Correlation	Wilks' Lambda	Chi-squared	df	Level of Significance
1*[a]	3.24275	53.93	53.93	.874	.0380	256.63	144	.0000
2*	1.70675	28.38	82.31	.794	.1614	143.19	105	.0079
3	.77066	12.82	95.13	.660	.4368	65.020	68	.5800
4	.29294	4.87	100.00	.476	.7734	20.168	33	.9611

Note. [a]Four functions can be generated from five groups.

Table 12 lists the standarized canonical discriminant function coefficients for each one of the behavioral criteria.

Table 12

Standardized Canonical Discriminant Function Coefficients for the First Two Significant Functions

	Dimension Criteria	Function 1 Coefficients	Function 2 Coefficients
P	Effective Preparation	-.837	-.241
P1	Perceives the task goal as a problem solving situation	.397	.275
P2	Identifies own position in relation to task goal	.404	-.431
P3	Describes the key dimensions of the interaction setting	-.217	.022
P4	Develops a chronological strategy for assisting the group	.501	.311
P5	Clearly states the purpose of the task	-.006	.207
P6	Relates theory to practice	-.168	.325
P7	Cites information from outside sources	-.058	-.024
P8	Synthesizes the thinking of others	-.149	-.688
D	Effective Demonstration	-1.292	-.126
D1	Clearly states the purpose of the task	.339	-.291
D2	Review the task goal whenever appropriate	.267	-.191
D3	Summarizes or evaluates group progress	.203	.465
D4	Stimulates discussion through effective questioning	-.174	-.648
D5	Shows enthusiasm for goal attainment - nonverbal behavior	-.074	-.467
D6	Assigns tasks to help implement decisions	-.642	-.150
D7	Maintains appropriate eye contact	-.529	-.130
D8	Uses both person-oriented and task-oriented skills	.230	.462
D9	Uses a broad spectrum of interaction skills	.552	-.627

Table 12 (continued

	Dimension Criteria	Function 1 Coefficients	Function 2 Coefficients
D10	Chooses interaction skills appropriate to the situation	.240	-.260
D11	Attends to non-verbal messages	.022	.502
D12	Encourages those who exhibit withdrawing behavior	.070	.578
D13	Challenges those who exhibit interfering behavior	.141	.252
D14	Mediates differences	-.528	-.182
D15	Listens to others without interrupting when challenged	-.437	-.315
D16	Maintains an attitude of openness	.799	.462
D17	Shows ability to modify ideas when appropriate	.115	-.213
D18	Contributes a fair share of ideas	.447	-.213
D19	Points out implications of ideas presented	.160	.674
S	Effective Self-Assessment	-.389	.820
S1	Shows how the demonstrated behavior had an effect	-.092	-.004
S2	Provides an adequate rationale for any changes made in plan	.229	-.235
S3	Cites examples of effective behaviors and states how they affected others	.627	-.297
S4	Identifies strengths and weaknesses of own performance	.740	-.785
S5	Points out ways to improve one's effectiveness	.338	-.572
L	Effective Leadership	.395	.677

For each function, the behavioral criteria were rank ordered from high positive to high negative. The criteria which contribute high positive values describe the group having the higher positive mean score on that function. The negative values are more descriptive of the group which has a negative mean score on that function.

The four highest positive values or coefficients and the four highest negative coefficients were selected out from function 1 (Table 12).[1]

High positive criteria on function 1:

D16	Maintaining attitude of openness	.799
S4	Identifies strengths and weaknesses of own performance	.740
S3	Cites examples of effective behaviors	.627
D9	Uses a broad spectrum of interaction skills	.552
P4	Develops chronological strategy	.501

High negative criteria on function 1:

D	Effective demonstration	-1.292
P	Effective preparation	-.837
D6	Assigns tasks	-.642
D14	Mediates differences	-.528
D7	Maintains eye contact	-.529

High positive criteria on function 2:

S	Effective self-assessment	.820
L	Effective leadership	.677
D19	Points out implications of ideas	.674
D12	Encourages those who exhibit withdrawing behavior	.578
D11	Attends to non-verbal messages	.502

High negative criteria on function 2:

S4	Identifies strengths and weaknesses of own performance	-.785
P8	Synthesizes the thinking of others	-.688
D4	Stimulates discussion	-.648
D9	Uses a broad spectrum of interaction skills	-.627
S5	Points out ways to improve own effectiveness	-.572

[1]Those coefficients selected "stand out" from the rest, e.g., they may be twice or three times as high.

The means of the two discriminant functions for the five groups were then computed:

	Function 1	Function 2
Group I (WEC, uninstructed, older, prior college)	.430	-1.614
Group II (WEC, instructed, older, no prior college)	1.316	- .339
Group III (WEC, uninstructed, younger, no prior college)	1.554	1.218
Group IV (WDC, instructed, younger, no prior college, no contract)	-1.162	1.621
Group V (WDC, instructed, younger, no prior college, contract)	-4.335	- .915

A plot of the means (centroids) is shown in Figure 7. This figure shows that the first discriminant function separates the three Weekend College Groups I, II and III from the Weekday College Groups, IV and V. The contracted Group V is further from all the groups. Function 1 accounts for 54% of the variance; thus we have more confidence in function 1 criteria as describing the groups.

Figure 7 also shows the means with respect to function 2. The second discriminant function separates the two groups who are younger (III and IV) from Groups I, II and V. Function 2 accounts for 28% of the variance.

A student who scored high positive on the first function is characterized by openness, ability to reflect on her own experience in citing effective behaviors, using a broad spectrum of interaction skills, and effectiveness in preparing by developing a chronological strategy. This profile of behaviors is more characteristic of the Weekend College Groups I, II and III who scored high positive on function 1.

A student who scored high negative on the first function is characterized by effective Demonstration, effective Preparation, assigning tasks, mediating differences, and maintaining eye contact. These criteria are more characteristic of the Weekday College groups. The contracted group in particular scored high negative on function 1.

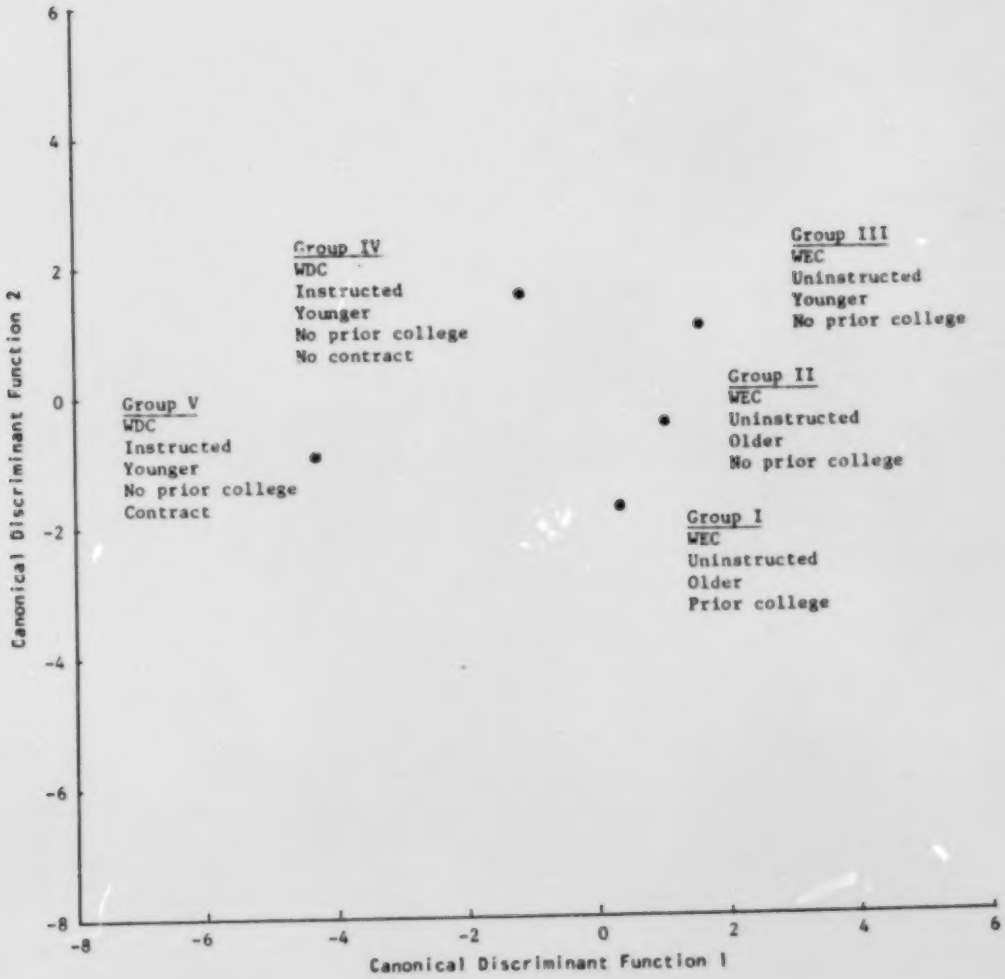

Figure 7. Group Centroids With Respect to Function 1 and Function 2.

Prior to data analysis, groups were categorized on the basis of program (WEC/WDC). The discriminant analysis reinforces this categorization. With respect to function 1, Weekend College groups highly resemble each other and Weekday groups highly resemble each other.

Now let us concern ourselves with function 2. A person who scores high positive on function 2 is characterized by effective Self-Assessment, effective Leadership, pointing out implications of ideas, and attentiveness to group members (attends to nonverbal messages, encourages those who exhibit withdrawing). These behaviors best describe Groups III (WEC, uninstructed, younger, no prior college) and IV (WDC, instructed, younger, no prior college, no contract) which are the two younger student groups. These younger groups, from both Weekday and Weekend College, differ on instruction, however.

A person who scores high negative on function 2 will synthesize thinking of others, stimulate discussion, use a broad spectrum of interaction skills, identify her strengths and weaknesses and point out ways to improve her own effectiveness. This profile of behaviors characterizes Group V (WDC, instructed, younger, no prior college, contract), Group I (WEC, uninstructed, older, prior college), Group II (WEC, uninstructed, older, no prior college) and Group IV (WDC, instructed, younger, no prior college, no contract).

The classification analysis also shows how students within groups are similar. Table 13 supports our categorization by depicting intragroup membership stability. The discriminant analysis confirmed that Weekday College students (Groups IV and V) are more similar among themselves than they are to individuals in the other three Weekend College groups (I, II, III). The highest prediction for correct group membership

Table 13

Percentage of Correct Classification for Each Group

Actual Group Membership	n	Predicted Group Membership									
		I		II		III		IV		V	
		n	%	n	%	n	%	n	%	n	%
Group I WEC Uninstructed Older Prior college	25	20	80%	3	12%	0	0%	1	4%	1	4%
Group II WEC Uninstructed Older No prior college	25	3	12%	19	76%	3	12%	0	0%	0	0%
Group III WEC Uninstructed Younger No prior college	17	0	0%	3	17.6%	14	82.4%	0	0%	0	0%
Group IV WDC Instructed Younger No prior college No contract	23	0	0%	2	8.7%	0	0%	21	91.3%	0	0%
Group V WDC Instructed Younger No prior college Contract	10	0	0%	0	0%	0	0%	0	0%	10	100.0%

classification occurred in Group IV (91% correctly classified) and Group V (100% correctly classified).

The group categorization prior to our first analysis was made based on differences in age and other college experience for the uninstructed Weekend College students (Groups I, II, III). The instructed Weekday College students were also divided according to contract/no-contract.

In the previous univariate analysis, age and prior college experience did not emerge as factors affecting Social Interaction performance. The classification analysis shows, however, that intra-group similarities were higher than inter-group similarities. This confirms our earlier classifications.

Leadership Dimension Score Analysis

To further investigate the meaning of the Social Interaction competence, the Leadership dimension score was examined in relationship to the behavioral criteria.

- What relationships exist between Leadership and type of program, instruction, age and other college experience?

Correlation matrices for each group were generated to identify behavioral criteria which correlate highly with the Leadership dimension score within each group, and to identify which of these highly correlated behaviors hold constant across groups (see Table 14).

The behaviors which yielded high inter-correlations ($p < .01$) with Leadership scores irrespective of group membership are:

D4 Stimulates discussion through effective goal related questioning

D10 Chooses interaction skills appropriate to the situation

D18 Contributes a fair share of ideas and suggestions

Group V (WDC, instructed, younger, no prior college, contract) had very few behaviors which correlate highly with Leadership scores (8 out of 32). The nature of the exercise in which students are credentialed for level 4 validation may have affected the frequency occurrence of Leadership behaviors. The other groups had an average of 20 behavioral criteria that are highly correlated with the Leadership dimension score. Table 14 shows the list of criteria which are highly correlated with the Leadership dimension score.

As mentioned earlier, Group V (WDC, instructed, younger, no prior college, contract) students were expected to demonstrate more self-initiated behaviors because they were contracted.

If Group V (WDC, instructed, younger, no prior college, contract) is discarded from the Leadership criteria correlation matrix, other behaviors may also be added to the list of behavioral criteria which highly correlate with Leadership. These Leadership criteria relationships hold across groups.

The additional behaviors are:

P6 Relates theory to practice

D1 Clearly states the purpose of the task

D3 Summarizes or evaluates group progress in relation to goal achievement

D6 Assigns tasks to help implement decisions or attain consensus

D9 Uses a broad spectrum of interaction skills

D12 Encourages those who exhibit withdrawing behavior

A Leadership dimension emerges from this correlational analysis. Since some Leadership criteria hold constant across all groups (excluding Group V), we may conclude that these behavioral criteria describing Leadership are generic rather than situation or group specific. The

Table 14

Behavioral Criteria Within Groups Which Were Significantly Correlated with the Leadership Dimension Score

Group I	Group II	Group III	Group IV	Group V
WEC	WEC	WEC	WDC	WDC
Uninstructed	Uninstructed	Uninstructed	Instructed	Instructed
Older	Older	Younger	Younger	Younger
Prior college	No prior college	No prior college	No prior college	No prior college
			No contract	Contract
P	P_1	P_3	P_6	P_5
P_4	P_3	P_5	D_1	D_4
P_6	P_4	P_6	D_2	D_{10}
P_7	P_6	D_1	D_3	D_{18}
D_1	P_7	D_2	D_4	D_{19}
D_3	D_1	D_3	D_5	S_1
D_4	D_2	D_4	D_6	S_4
D_5	D_3	D_5	D_7	S_5
D_6	D_4	D_6	D_9	
D_7	D_6	D_7	D_{10}	
D_8	D_8	D_8	D_{12}	
D_9	D_9	D_9	D_{16}	
D_{10}	D_{10}	D_{10}	D_{18}	
D_{11}	D_{11}	D_{12}	D_{19}	
D_{12}	D_{12}	D_{14}	S_3	
D_{14}	D_{15}	D_{18}	S_4	
D_{16}	D_{16}	D_{19}		
D_{17}	D_{17}	S_1		
D_{18}	D_{18}	S_3		
D_{19}	S_1			
S_3	S_2			
	S_4			

general description of a leader thus includes person- and task-oriented skills. A student demonstrating person-oriented behaviors is sensitive to group members' needs, and uses a wide range of interaction skills appropriately. A student demonstrating task-oriented skills stimulates discussion, contributes ideas, cites information summarized and assigns tasks. She takes on an active role in the group discussion.

CONCLUSIONS

Instructional Validity

Findings from the group comparisons indicated that instructed and uninstructed students who completed the Integrated Competence Seminar perform similarly on each one of the Social Interaction dimensions: Preparation, Demonstration, and Self-Assessment. Some significant differences were found on the Preparation and Self-Assessment dimensions. The directions of the differences, however, were not consistent enough to support the instructional validity of the Social Interaction competence. No differences were obtained on the Demonstration dimension, which represents the performance aspect of the Social Interaction competence.

Effective performance on Preparation is reflected in the degree to which a student prepares to present her ideas during the Group Discussion. Apparently, the motivation to achieve a level 4 validation enhanced the performance of the contracted group on the Preparation dimension.

One of the younger uninstructed groups was most effective in identifying strengths and weaknesses in their own performance and pointing out ways to improve Social Interaction effectiveness. No relationships were found between effective Preparation and effective Demonstration or between effective Self-Assessment and effective Demonstration. It is assumed, however, that learning activities which include greater emphasis on Preparation and Self-Assessment skills may assist students in their Demonstration performance, thus assisting students to plan strategies prior to group interaction, and to direct their attention to assessing their performance after completing the task.

The ANOVA findings did not yield group main effects. These results support our previous finding of similarity of performance among groups, and no consistent direction of higher performance by the instructed group. The dimension main effect shows intra-group differences rather than inter-group differences on the various dimensions.

There are some possible explanations for these findings. One centers on the validity of the Integrated Competence Seminar as an adequate measure of level 4 performance, because of the relationship of the Integrated Competence Seminar (ICS) assessment technique to Social Interaction learning experiences. During the first two years of her education, a student is expected to accomplish one validation at level 1, three validations at level 2, one validation at level 3, and five validations at level 4 in order to complete her Social Interaction requirements. The student is credentialed within various classes which offer Social Interaction validations. The ICS assessment procedure for most students is their first open-ended, uncredentialed, unstructured experience. The student is expected to transfer her Social Interaction abilities to the ICS "realistic situation." This is her first attempt to translate her acquired abilities into an out-of-class experience and test the "outcomes" of her learning experiences. Thus far, she has been involved in the process of acquiring the desired behaviors in class. Now she is expected to demonstrate her ability and integrate previous learnings in an unknown situation. She is asked to move from "guided instruction" to a self-initiated activity. In that respect, she is <u>more similar than different</u> from the incoming students who are also encountering an open-ended task-oriented group discussion for the first time.

The study pointed out that at this juncture instructed students do not differ significantly from uninstructed students. Since most instructed students (Weekday) achieved their Social Interaction validation requirements in classes prior to their ICS assessment, one may postulate that the problem of lack of instructional validity lies in the issue of transfer. If faculty have evidence that social interaction skills learned in classes are effective and sufficient in enhancing Social Interaction performance within any given situation, the problem of adequate learning activities for successful transfer should be explored. Several approaches could be considered:

- Implement a variety of simulation exercises of an open-ended nature throughout the four years in college.
- Emphasize open-ended simulation exercises at the advanced levels of Social Interaction activities (e.g., mainly during the last two years).
- Prepare the students for ICS-type activities by reducing structured learning in class from the beginning.

Comparing instructed and uninstructed students on a "real life" exercise after only two years in college may be a premature attempt to demonstrate instructional validity. The question here may not be instructional validity but rather validity of transfer, especially with the Social Interaction competence, where uninstructed students encounter similar situations almost daily. If we assume that Alverno students are provided with the necessary tools to interact effectively, an additional step is needed to ensure that these tools can be brought in a variety of out-of-class situations.

Another observation has to do with the extent to which the Group Discussion in the Integrated Competence Seminar elicited higher level cognitive performance and content integration. The fact that the Weekday College instructed students performed as well as more mature, uninstructed students in the Weekend College may show that Social Interaction can be taught. No differences between the groups is a positive result. But it may also be the case that it is not until levels 5 and 6 that differences between uninstructed and instructed students show up. Still, the Social Interaction Division may need to be more selective in identifying criteria at level 4, criteria that expect higher level cognitive performance and content integration. The integration of content and competence, which is not measured by the "content-fair" Integrated Competence Seminar, and the ability to analyze behaviors and group behaviors may be the difference between uninstructed and instructed students. Given these results, it may be that the Preparation dimension is the key to the difference. Preparation and Self-Assessment skills may develop first.

Construct Validity

As stated earlier in the design section, different Social Interaction performance patterns may imply that different constructs are employed by students from different groups in producing desired behaviors. Patterns of behaviors which hold constant across dissimilar groups may indicate a generic construct. The question of construct validity as it is related to different groups, instructed vs. uninstructed, as well as different backgrounds, is particularly important when one views construct validity as focusing on response and interpretation of scores (Messick, 1975),

where the validity of interpretation is the main emphasis rather than the validity of the instrument.

As Nitko (1980) states: "Following the cognitive-psychology views summarized by Chi and Glaser (in press) one may find, for example, that experts and masters exhibit different behaviors and may use different internal cognitive processes when attempting to answer the same test items. Measuring these kinds of differences sharply contrasts with a measurement model that is linked to a conception of mastery that is operationalized as a certain proportion of correct answers on a well-defined pool of items."

A study design which incorporates five dissimilar groups described by a wide range of different variables may assist faculty in understanding the differences as well as similarities in Social Interaction behaviors which characterize each group. Can we identify patterns of Social Interaction behavior which characterize each group of students? Does the Social Interaction construct as a defined set of abilities have the same meaning for different groups, or rather, are different groups identified by different Social Interaction skills?

Since the univariate analysis did not yield major differences among pairs of groups, it is particularly interesting to see what kind of differences emerge in a multivariate analysis which takes the inter-correlations among the variables into account. The means of the five groups with respect to the first discriminant function, which was highly significant and contributed 53 percent to the separation among the groups, indicated (in spite of the previous findings) that the three uninstructed groups exhibited closer proximity on the multivariate space while the two instructed groups were also close in the opposite direction. That may imply similar underlying constructs employed by three uninstructed

groups, and similar underlying constructs employed by the instructed groups. Since the major difference between the two groups was the instructional factor, one may postulate that behaviors which discriminate the uninstructed groups are descriptive of baseline behaviors prior to instruction, whereas behaviors which discriminate the instructed groups are descriptive of performance following instruction. Since no differences were found on the Demonstration dimension--the performance aspect of the assessment procedure--we looked closely at the Demonstration behaviors which appear to separate the uninstructed and the instructed groups, with respect to the two functions.

The first function indicated that the uninstructed entering groups are described by general Social Interaction behaviors such as maintaining an attitude of openness, or using a broad spectrum of interaction skills. When the instructed profile is examined, task-oriented skills as well as person-oriented skills are evident: assigns tasks (task), maintains eye contact (person), indicates differences (person), or stimulates discussion through questioning (task). Thus, the effects of instruction as implied by these behaviors are task- and person-oriented behaviors which are employed by instructed students for effective group interaction skills.

In the second significant function, the best comparison groups, controlling for experience, resembled each other and are closer on the multivariate space. The common factor for these two groups is traditional college age. This may indicate a maturity factor which interacts with social interaction performance.

The two younger groups are characterized by person-oriented behavior: attends to non-verbal messages, or encourages those who exhibit withdrawing behavior. Are these behaviors descriptive of younger students?

The older uninstructed groups and the contracted instructed group are all clustered toward the negative dimension with respect to the second function. A factor that may explain this cluster is motivation since the uninstructed groups were older and perhaps more mature and the younger group was contracted. The dominant factor might be age if the 10 Weekday instructed students are excluded. If we exclude these 10 students, older Weekend College women are characterized by broad interaction skills and an ability to stimulate discussion through questioning. This skill was identified as an outcome of instruction according to the first function. Since it emerged in the uninstructed group as well, attributing it to the effects of instruction may be premature. But the fact that young, instructed students perform this ability as well as mature women, given the nature of the ability under study, argues for the validity of instruction.

Since on both functions, older, uninstructed women are described as having broad interaction skills, and since broad interaction skills signify level 1 of the Social Interaction competence, it is reasonable to assume that women entering Weekend College may need less extensive instruction on Social Interaction level 1 behaviors. It is suggested that faculty examine the possibility of shifting instructional efforts toward the higher levels of Social Interaction for these groups.

The faculty member who scored the generic instruments for this study used information about level 1 performance from the external assessors' evaluation of student performance at level 1. The fact that uninstructed students demonstrated high performance on level 1 behaviors according to external assessors' judgment, implies that external assessors are validating students quite readily. This issue of lower standards at level 1 for entering students should be examined

as well as why instructed students did not perform higher on level 1 behaviors.

When Leadership characteristics are examined across all five groups, the analysis yielded a list of leadership behaviors which hold constant across all groups and could be considered a generic profile of leadership ability, which includes person- and task-oriented skills. Thus, an emerging leader in a group could not be effective in focusing only on task-oriented behaviors and ignoring person-oriented behaviors, nor can she focus only on interpersonal behaviors and ignore task-oriented behaviors. This supports the Social Interaction Division's recent incorporation of a person- as well as a task-oriented model. A balanced approach to both person- and task-oriented behaviors will more likely produce an effective leader.

The contracted, instructed group showed a marked reduction in leadership behaviors. It was assumed that performance under the pressure of validation produced a conflict of interest--achievement of group task vs. effective individual performance for credentialing purposes. Under such circumstances, all students are trying to exhibit as many behaviors as possible. There may be no occasion to demonstrate other leadership behaviors which are other-oriented in nature. This argues for an assessment technique that elicits more collaboration.

In conclusion, the discriminant analysis assisted in establishing construct characteristics of the major two groups: instructed vs. uninstructed. It showed clearly that uninstructed students interpret effective Social Interaction skills differently from instructed students. If the instructed group description supports the instructional framework, then faculty may develop an instructional program which will take into

account the baseline behaviors of entering students and further develop the desired behaviors which did not describe the instructed groups.

Criteria Validity

This section of the analysis dealt with criteria evaluation. Can we identify criteria which best discriminate instructed and uninstructed groups? A combined group comparison where the total instructed group response was compared to the total uninstructed groups' response yielded three discriminative criteria:

- Assigns tasks to help implement decisions or attain consensus
- Maintains appropriate eye contact with everyone in the group
- Attends to non-verbal messages and cues

A paired group comparison yielded an additional criterion which appeared to discriminate between the two groups providing the most accurate comparison for instructed/uninstructed.

- Stimulates discussion through effective goal related questioning

All uninstructed groups performed consistently higher on:

- Develops a chronological strategy for assisting the group toward goal achievement
- Identifies strengths and weaknesses of own performance
- Points out ways to improve one's own effectiveness.

The faculty coder felt that the forms uninstructed students filled out prior to Group Discussion (Preparation dimension) were not comparable to the forms completed by instructed students. The former stimulus elicited "a chronological strategy for assisting the group towards goal achievement" more explicitly. Thus, it cannot be concluded that uninstructed students are more effective in developing chronological

strategies. Finally, two other Self-Assessment criteria yielded higher performance in the uninstructed groups; this phenomenon should be further investigated as a step toward identifying levels of self-assessment ability.

Overall, the generic instrument for Social Interaction level 4 did not indicate high discriminative power between instructed and uninstructed students. Since the generic instrument was employed in a special assessment procedure--the Integrated Competence Seminar--one has to determine if transfer of abilities actually occurred from classroom learning to the ICS simulation exercise before instrument validity can be questioned. As stated earlier, the generic instrument is an attempt to create a common measure across all Social Interaction assessment procedures at level 4. However, since the ICS Group Discussion offers a different assessment procedure from what is offered regularly in classes, adapting this measure which was generated from classroom activities to a new open-ended situation is questionable. Does the ICS Group Discussion elicit the behaviors which are measured by the generic instrument? These questions represent basic issues of the instrument relationship to the assessment mode or to the assessment stimuli.

A multiple regression analysis was employed to assist faculty in identifying criteria which contribute to or may improve the prediction of the dimension score: effective Preparation, effective Demonstration, effective Self-Assessment. This type of analysis will assist the Social Interaction Division to zero in on the "important" behaviors which underlie effective Preparation, effective Demonstration, and effective Self-Assessment. Faculty then can qualitatively evaluate the remaining criteria and their relationship to the dimension construct and also identify baseline performance of the behaviors. It will also assist

in identifying a more demanding set of criteria at level 4, which insist on the integration of content and competence.

DISCUSSION

This study had several major outcomes for our work in validating assessment techniques. First, it led us to compare and contrast several statistical methods for analyzing the results from an attempt to validate still another measure. Second, the findings have several implications for validation of curricula in general. Third, we came to better understand the nature of the Social Interaction competence, leading to the validation of the competence as well as the technique.

The first finding is that group comparisons between traditional-aged instructed and mature uninstructed students may not be an effective strategy for validating assessment techniques if the competence is one developed through informal as well as college learning. Such competences may also interact with level of ego development. And such a strategy may be premature where little is understood about the nature of the ability. In general, the study indicated similarities in performance between instructed and uninstructed students in a pre-/posttest design. When we ignored the extent to which students had been instructed however, and combined both groups in a discriminant analysis, some clear findings emerged. We found that some of the Social Interaction abilities did indeed separate the uninstructed students as a group from the instructed students. The criteria that discriminate the instructed students do show evidence of the effects of instruction, in that they are more closely related to those aspects of Social Interaction that are learned as part of a more specific social interaction process.

Realizing the possible effects on performance of such variables as age, type of program, and prior college experience, we tried to control for these variables at first. Such an undertaking, given the nature of the competence we were assessing, was in retrospect doomed to failure.

Despite our efforts, we were simply not able to control for the myriad range of variables that are likely to affect the results, and the beginning section of this paper is a litany of description of all the other possible aspects of the sample, the conditions of the assessment, and the assessment process that were at least partially uncontrolled. When this is combined with the nature of the competence under study, where performance also interacts with a set of personal and ego development variables, trying to separate out the specific effects of instruction that show significant differences between small sample sizes through group comparison is not an effective strategy.

Group comparisons, where persons are difficult to categorize and matching is impractical because of small sample sizes, may not be the best strategy for evaluating the effects of instruction. One outcome of the analyses, then, is that a strong focus on group comparison, where the groups are independent rather than longitudinal, may not be an effective strategy unless the groups can be matched. Another goal, to include a broad range of students on age and experience, did lead to an effective strategy for identifying those Social Interaction criteria that validate the construct. Getting at the construct is the most important task. The effects of instruction are also important, but in Social Interaction, we do not yet know enough about the integration of age and experience with performance. Construct validity should take priority in such cases. This is probably best accomplished through multivariate rather than univariate techniques.

In the main, our analysis shows that the study of assessment techniques should not be limited to univariate methods. Even if differences do not appear in group comparisons, a great deal can be learned about the construct under study, and the effects of instruction. Compared to

univariate analyses, multivariate analyses give us a more holistic view of the performance under study. It can assist us in identifying person categories, and patterns of coherent group performance that emerge in the data. This also provides us with a more holistic picture of group characteristics, whereas a univariate analysis can look at only one factor at a time. This is particularly important in relation to construct validity since we wish to examine the similarities and differences in performance that describe the different groups. From the regression analyses, we can examine the kinds of constructs assessors are using to evaluate students. Such analyses give us a glimpse of the criteria an assessor emphasized in describing student performance.

In education, "clean" studies are seldom possible given limitations such as those described in the design section. Matched samples or longitudinal pre-/posttest designs may be impossible. Putting all students into the same analysis allows us to see which groupings occur in the data, even though the results from the discriminant analysis must be interpreted with caution and results used as indicators rather than conclusions.

In our prior study of the Valuing and Communications competences, we adopted the idea of studying a wide range of students in order to investigate construct validity. In the present study, this strategy provided the most useful information for both construct validity and instructional design purposes.

With respect to an overall design, we do recommend, for future studies,a longitudinal pre-/posttest design with a particular class or from one competence level unit to another. While a longitudinal study over two years may not work because of the continual redesign of

curriculum and assessment techniques, we might follow one group from level 1 to level 2, another from level 2 to level 3, another group from level 3 to level 4, and so on. Pre-/posttest designs could be used better as a particular strategy for studying changes in student performance within courses, even though effects may not necessarily be attributed to specific aspects of instruction in a particular course because of the lack of a control group.

The issue here is the value of two group comparisons. At this level, it has been ineffective. By the time we specify all the qualifications to the results we begin to wonder why we are doing the study. Clearly, these comparisons need to control variables. But the intensive study of which variables contribute to the variance in performance takes considerable time and effort. Such study is best left to external validation efforts, rather than to internal efforts, in our opinion. Further, faculty need to be able to predict performance from those variables, but at the individual student level. An interview by an experienced faculty member can quickly elicit information on prior college experience, and other levels of understanding, and the instructor may then use the information to prescribe instruction. Identifying qualifications for explaining results is a good exercise, however. It reinforces one criterion for the assessment process, the need for multiple samples of behavior.

Methodology chosen for validation needs to be set within the context of where the instrument is in the validation process. At the same time, one cannot wait until an instrument is "perfect" before one initiates studies of validity. We need an ongoing process, and sometimes, given the necessity of information for the validity of the construct of the competence, we will validate an instrument for that purpose.

Clearly, this study supports the need for our inplace, extensive evaluation/revision system (Assessment Committee/Office of Research and Evaluation, 1980) which combines a variety of strategies and depends particularly on the level of validity of the assessment technique. This system recognizes that all instruments are always in process. At any one point in time, an instrument may have a particular level of validity. For example, it may have been reviewed by one member of the department and used one semester so there is performance data available. At another level, it may have been reviewed by the Assessment Council against the criteria for judging instruments, and it may have had a criteria evaluation. At still another level, inter-rater reliability of assessor judgments may have been established, and student performance analyzed for input into criteria for student credentialing.

In sum, an important methodological strategy for beginning work on validation of an instrument measuring a competence that is not well defined, is to collect information from a wide range of students who differ on level of maturity, experience and instruction and examine their performance to identify those criteria which discriminate the categorization of individuals. This will give us a better idea of the holistic nature of performance on the competence. Another benefit is that we will then be able to identify the criteria that are most likely to provide faculty with indicators that a student is mastering the ability. One problem in designing criteria is choosing those criteria from a group which give the best picture of the ability under study. Such a discriminant analysis can assist with this effort.

In conclusion, the present study outlined a procedure by which the integration of information about a competence construct, different group characteristics and their relationship to behavioral criteria, and criteria

evaluation contribute to an information base for instructional development, re-evaluation of competence behavioral criteria and revision of instruments which measure these behaviors.

REFERENCES

Assessment Committee/Office of Research and Evaluation. Validating assessment techniques in an outcome-centered liberal arts curriculum: Insights from the evaluation and revision process. Milwaukee, WI: Alverno Productions, 1980.

Assessment Committee/Office of Research and Evaluation. Validating assessment techniques in an outcome-centered liberal arts curriculum: Integrated Competence Seminar. Milwaukee, WI: Alverno Productions, 1982.

Assessment Committee/Office of Research and Evaluation. Validating assessment techniques in an outcome-centered liberal arts curriculum: Six Performance Characteristics rating. Milwaukee, WI: Alverno Productions, 1983.

Friedman, M., Mentkowski, M., Earley, M., Loacker, G., & Diez, M. Validating assessment techniques in an outcome-centered liberal arts curriculum: Valuing and communications generic instruments. Paper presented at the meeting of the American Educational Research Association, Boston, April 1980.

Glass, G. & Stanley, J. Statistical methods in education and psychology. New York: Prentice-Hall, 1970.

Klecka, W. Discriminant analysis. In Nie, N., Hull, C., Jenkins, J., Steinbrenner, K. & Brent, D. Statistical package for the social sciences, 2nd ed. New York: McGraw-Hill Book Company, 1975.

Mentkowski, M., & Doherty, A. Careering after college: Establishing the validity of the abilities learned in college for later success (NIE-G-77-0058). Milwaukee, WI: Alverno College, 1977.

Mentkowski, M., & Doherty, A. Careering after college: Establishing the validity of abilities learned in college for later careering and professional performance. Final report to the National Institute of Education. Milwaukee, WI: Alverno Productions, 1983.

Messick, S. Constructs and their vicissitudes in educational and psychological measurement. Psychological Bulletin, 1981, 89 (3) 575-588.

Nitko, A. Individual differences between criterion-referenced tests. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

3401 South 39th Street / Milwaukee, WI 53215